

CUESTIONES ÉTICAS



UN GRAN PODER

**CONLLEVA UNA GRAN
RESPONSABILIDAD**



→ CIENCIA Y
TECNOLOGÍA

Inteligencia artificial: 'Es probable que asistamos a la vegetalización de la humanidad'

SOFTWARE

Estos son los trabajos que no desaparecerán a causa de la inteligencia artificial

Instalarán cámaras con inteligencia artificial en trece colegios de Sevilla

SOCIEDAD · Exclusivo suscriptores +

La inteligencia artificial en la educación: los alumnos ya no copian, le piden a la máquina que les haga los trabajos

**Un abogado usó ChatGPT en un juicio.
Ahora es él quien debe dar
explicaciones a un juez por incluir citas
falsas**

**La Fiscalía estudia si las
fotos de menores creadas
con inteligencia artificial en
Almendralejo son delito**

**Voces clonadas con IA: nueva
tecnología para un viejo fraude**

Cine

**Hollywood se prepara para un
atasco de películas bajo la
amenaza de la inteligencia
artificial**

'Deep Voice' Software Can Clone Anyone's Voice With Just 3.7 Seconds of Audio

Using snippets of voices, Baidu's 'Deep Voice' can generate new speech, accents, and tones.



with DEEPMIND STARCRRAFT TRIUMPH

Let There Be Sight: How Deep Learning Is Helping the Blind 'See'



Technology outpacing security measures

Facial Recognition | Features and Interviews

AI beats docs in cancer spotting

A new study provides a fresh example of machine learning as an important diagnostic tool, Fred Bragier reports.

AI Can Help in Predicting Cryptocurrency Value

AI | Markets | 10 min read



'Creative' AlphaZero leads way for chess computers and, maybe, science

Former chess world champion Garry Kasparov likes what he sees of computer that could be used to find cures for diseases



How an A.I. 'Cat-and-Mouse Game' Generates Believable Fake Photos

By David Weiss and Nathan Vahedi | Jan. 4, 2018



Stock Predictions Based On AI: Is the Market Truly Predictable?

By David Weiss and Nathan Vahedi | Jan. 4, 2018



Google's DeepMind acers protein folding

By Robert F. Service | Dec. 9, 2018, 12:55 PM

After Millions of Trials, These Simulated Humans Learned to Do Perfect Backflips and Cartwheels

By David Weiss and Nathan Vahedi | Jan. 4, 2018



Neural networks everywhere

New chip reduces neural networks' power consumption by up to 95 percent, making them practical for battery-powered devices.

By David Weiss and Nathan Vahedi | Jan. 4, 2018



Researchers introduce a deep learning method that converts mono audio recordings into 3D sounds using video scenes

By David Weiss and Nathan Vahedi | Jan. 4, 2018



e faces show how far AI image generation has come in just four years

By David Weiss and Nathan Vahedi | Jan. 4, 2018



Automation And Algorithms: De-Risking Manufacturing With Artificial Intelligence

By Sarah Goodfellow | Contributor | Dec. 10, 2018

Editor on the subject of artificial intelligence

TWEET THIS

The two key applications of AI in manufacturing are pricing and manufacturability feedback

CONTENIDOS FALSOS

OPENAI'S NEW MULTITALENTED AI WRITES, TRANSLATES, AND SLANDERS

A step forward in AI text-generation that also spells trouble

By **James Vincent** | Feb 14, 2019, 12:00pm EST

Howard, co-founder of Fast.AI agrees. "I've been trying to warn people about this for a while," he says. "We have the technology to totally fill Twitter, email, and the web up with reasonable-sounding, context-appropriate prose, which would drown out all other speech and be impossible to filter."

CONTENIDOS FALSOS

Modelos de PLN pueden usarse para:

- Personificar la voz de una persona
- Generar noticias falsas
- Engañar a una persona para que piense que está hablando con una persona y no un chatbot

Cuando se use PLN para crear contenidos siempre hay que informar a los usuarios



MANIPULACIÓN

- Los modelos de PLN pueden emplearse para manipular
 - Campañas de desinformación, *trolling*
 - Intentos de *fishing*
- Curiosamente también pueden usarse para lo contrario:
 - Detectar *fake news* o trolls

DEEPPFAKES



SOBREVIGILANCIA



ÉTICA

Tiene impacto real en personas:

- ¿Quién se puede beneficiar? ¿A quién podemos dañar?
- ¿Se puede dar un mal uso a nuestro trabajo?



LENGUAJE Y PERSONAS

- *“The common misconception is that language has to do with words and what they mean. It doesn't. It has to do with people and what they mean” [H. Clark & M. Schober (1992)]*

LENGUAJE Y PERSONAS

El lenguaje no existe en soledad, proviene de personas y se usa para comunicarse con otras personas

- Dichas personas pueden tener sesgos y estereotipos
- Afecta al estado de sus producciones lingüísticas



SESGOS

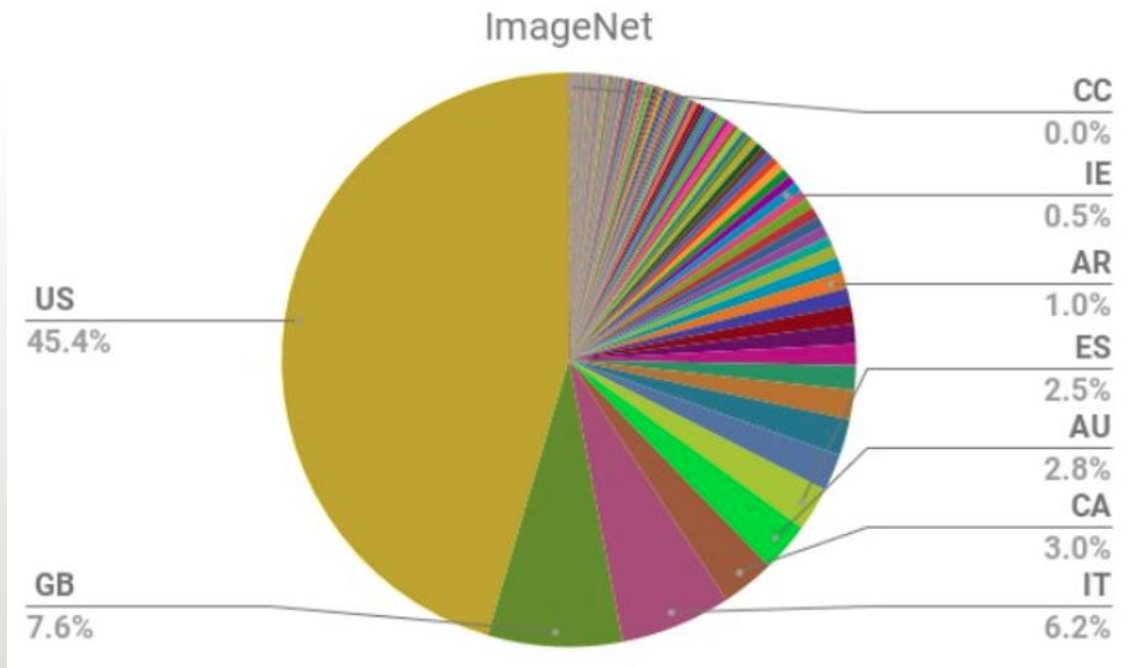
SESGOS

- 1.Sesgo de representación
- 2.Sesgo de medición
- 3.Sesgo de agregación
- 4.Sesgo de evaluación
- 5.Sesgo de despliegue
- 6.Sesgo histórico

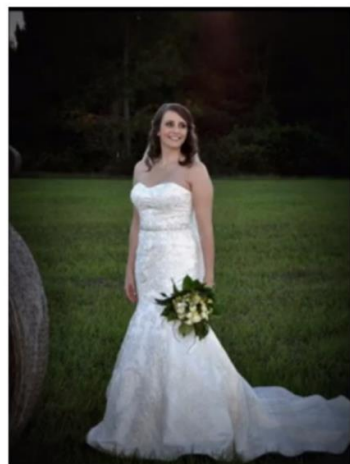
SESGOS DE REPRESENTACIÓN

El sesgo de representación se produce cuando los datos con lo que ha sido entrenado el modelo, infrarepresentan una parte de la población y como consecuencia, no es capaz de ofrecer una imagen amplia y diversa de la sociedad.

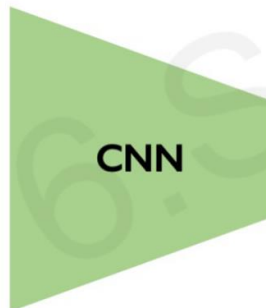
SESGOS



SESGOS DE REPRESENTACIÓN



Ground Truth: Bride



CNN for image
classification.



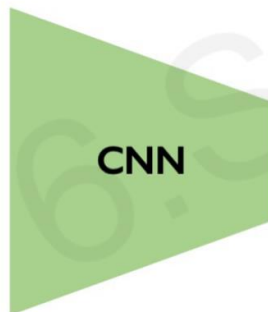
Predicted Classes

Bride ✓
Dress
Ceremony
Woman
Wedding

SESGOS DE REPRESENTACIÓN



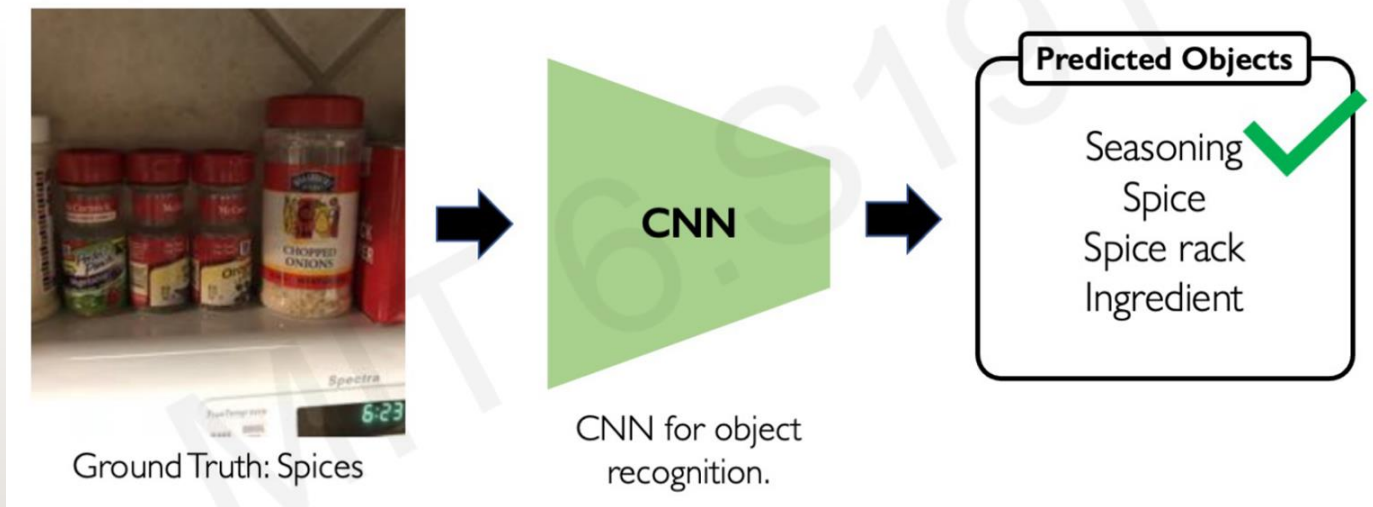
Ground Truth: Bride



CNN for image classification.



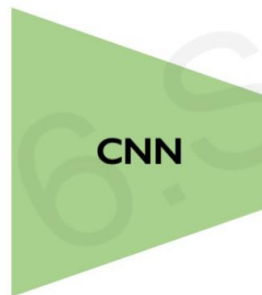
SESGOS DE REPRESENTACIÓN



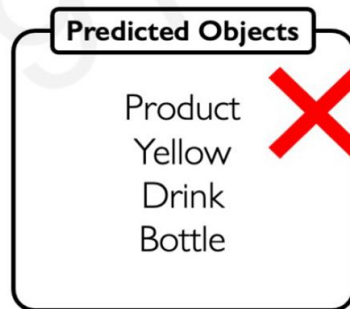
SESGOS DE REPRESENTACIÓN



Ground Truth: Spices



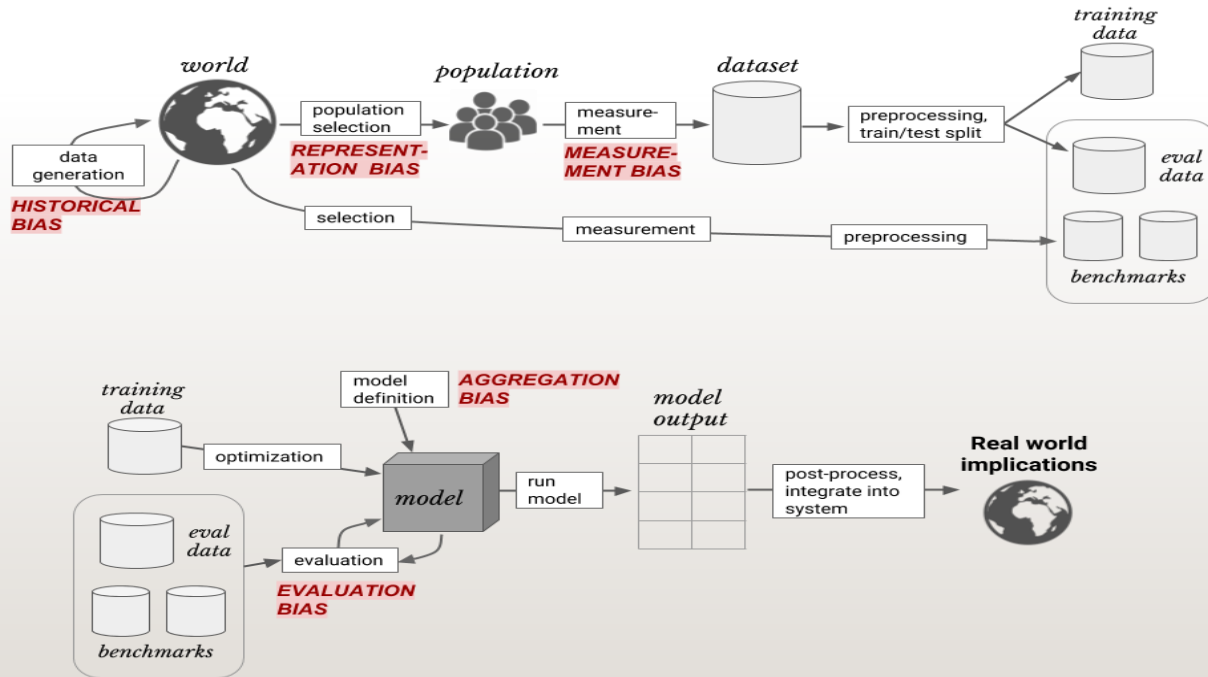
CNN for object recognition.



SESGOS DE MEDICIÓN

El sesgo de medición se produce al elegir, recopilar o calcular las características y etiquetas que se utilizarán en un proceso de la predicción a futuro. Normalmente, se produce cuando se quiere analizar una característica o idea que no es directamente observable.

SESGOS ALGORÍTMICOS



El algoritmo que ‘adivina’ los delitos futuros falla tanto como un humano

El sofisticado programa COMPAS analiza la posibilidad de reincidir de un millón de convictos reales

Machine Bias

There's software used across the country to predict future criminals. And it's biased against blacks.

by Julia Angwin, Jeff Larson, Surya Mattu and Lauren Kirchner, ProPublica

May 23, 2016

Prediction Fails Differently for Black Defendants

	WHITE	AFRICAN AMERICAN
Labeled Higher Risk, But Didn't Re-Offend	23.5%	44.9%
Labeled Lower Risk, Yet Did Re-Offend	47.7%	28.0%

SCIENCE

WHAT HAPPENS WHEN AN ALGORITHM CUTS YOUR HEALTH CARE

By [Colin Lecher](#) | [@colinlecher](#) | Mar 21, 2018, 9:00am EDT

Illustrations by [William Joel](#); Photography by [Amelia Holowaty Krales](#)

SESGOS DE MEDIDA

Usando datos históricos sobre pacientes para predecir infartos

- Infartos previos
- Enfermedades cardiovasculares
- Accidente por enfermedad
- Colonoscopia
- Sinusitis

No se miden los infartos

Se mide: tengo síntomas, voy al doctor, me hacen pruebas y recibo un diagnóstico de infarto

Does Machine Learning Automate Moral Hazard and Error?[†]

*By SENDHIL MULLAINATHAN AND ZIAD OBERMEYER**

SESGOS DE AGREGACIÓN

El sesgo de agregación surge cuando se utiliza un modelo único para datos en los que hay grupos subyacentes o tipos de datos que deberían considerarse de forma diferente. El sesgo de agregación se basa en la suposición de que la correspondencia entre las entradas y las etiquetas es coherente en todos los subconjuntos de datos cuando en realidad y a menudo, no es así.

SESGOS DE AGREGACIÓN

Ocurren al hacer suposiciones que afectan la definición del modelo

Ejemplo: Los protestantes son más propensos a suicidarse que los católicos (no se tuvo en cuenta los países en los que vivían los dos grupos y sus condiciones)

SESGOS DE EVALUACIÓN

Ocurre cuando el conjunto de test no representa la población final a la que se va a aplicar el modelo

SESGOS DE EVALUACIÓN

¿Cuáles son los síntomas de un infarto?

Heart Attack
WARNING SIGNS

85% of heart damage takes place within the first 2 hours of a heart attack

KNOW THE SYMPTOMS & CALL 9-1-1 IMMEDIATELY

MEN vs. WOMEN


SYMPTOMS THAT COME AND GO AND FINALLY BECOME CONSTANT AND SEVERE:

MEN	WOMEN
Shortness of breath	Unusual fatigue
Fatigue	Lightheadedness and fainting
Jaw Pain	Discomfort in one or both arms, neck, shoulder, jaw or stomach
Chest pressure, burning, aching or tightness	Chest pressure, squeezing pain in the center of their chest
Pain that travels down one or both arms	Upper abdominal pressure or discomfort
Anxiety	Feeling of Fullness
Back Pain	Nausea/Vomiting
Feeling of Fullness	Cold Sweat
Nausea	

Premier Health

© Premier Health 2017

SESGOS DE EVALUACIÓN

Gender Classifier	Darker Male	Darker Female	Lighter Male	Lighter Female	Largest Gap
 Microsoft	94.0% 	79.2% 	100% 	98.3% 	20.8% 
 FACE++	99.3% 	65.5% 	99.2% 	94.0% 	33.8% 
 IBM	88.0% 	65.3% 	99.7% 	92.9% 	34.4% 



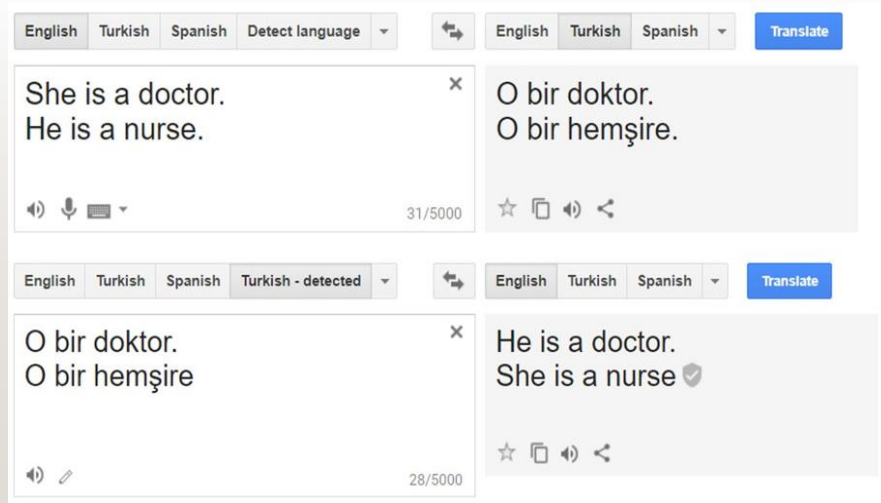
SESGOS DE DESPLIEGUE

El sesgo de despliegue o implantación se produce cuando no hay correspondencia entre el problema que se pretende resolver con un modelo y la forma en que se utiliza en la práctica

SESGOS HISTÓRICOS

Ocurre cuando el mundo actual no está alineado con los objetivos del modelo

SESGOS



SESGOS

The Rekognition Scan

Comparing input images to mugshot databases

1 INPUT: SEARCH IMAGES



2 RECOGNITION SEARCH



Mugshot Database

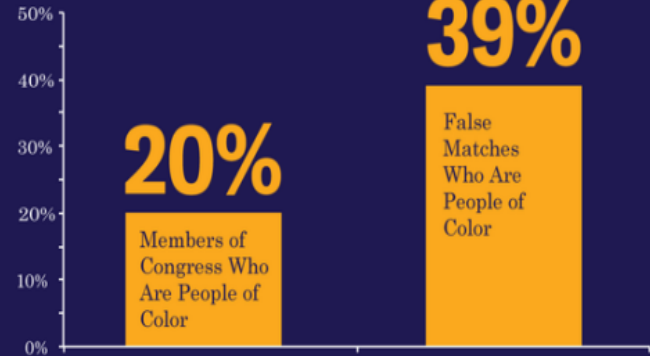
3 OUTPUT: PREDICTED MATCHES

NO MATCH

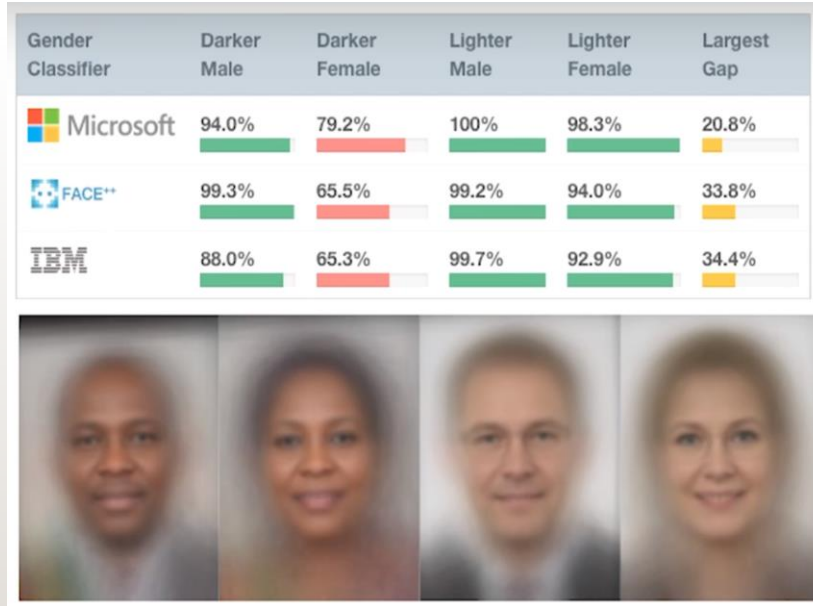
MATCH



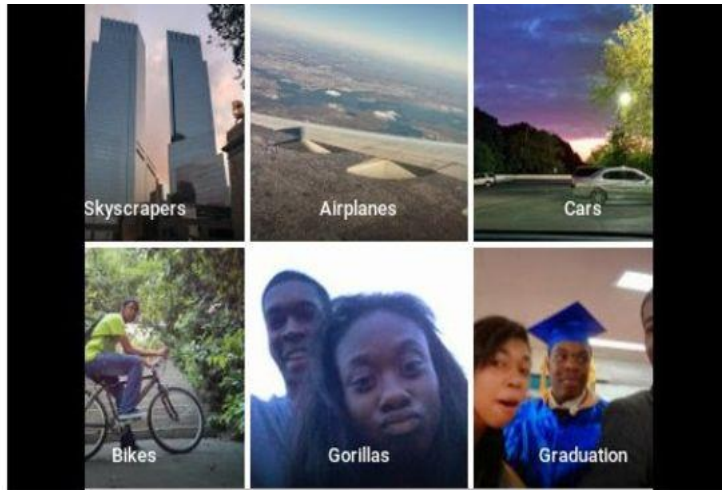
Racial Bias in Amazon Face Recognition



SESGOS



SESGOS



diri noir avec banan @jackyalcine · Jun 29
Google Photos, y'all [redacted] My friend's not a gorilla.



813



394



TWITTER

NEWS

Autonomous Cars Can't Recognise Pedestrians with Darker Skin Tones

People with darker skin more at risk of being hit by a self-driving vehicle.

SESGOS



SOLUCIONES

- Soluciones técnicas:
 - Usar datos menos sesgados y más inclusivos
- Transparencia y procesos de auditoría
- Fuerza de trabajo variada:
 - Crea aplicaciones menos sesgadas



Ley de IA de la UE: primera normativa sobre inteligencia artificial



https://www.europarl.europa.eu/pdfs/news/expert/2023/6/story/20230601STO93804/20230601STO93804_es.pdf

El uso de la inteligencia artificial en la UE estará regulado por la Ley de Inteligencia Artificial, la primera ley integral sobre IA del mundo. Aquí más información sobre cómo funcionará.

1. Medicina y atención médica
2. Automatización industrial
3. Conducción autónoma
4. Asistentes virtuales y chatbots
5. Finanzas y gestión de inversiones
6. Traducción automática
7. Agricultura de precisión

1. Educación personalizada
2. Detección de fraudes y ciberseguridad
3. Predicción del clima y desastres naturales
4. Música y entretenimiento
5. Ciencia y investigación
6. Cuidado de la salud mental
7. Control de calidad en la fabricación
8. Gestión de la cadena de suministro



AP photo

Elementary school teachers picket against use of calculators in grade school
The teachers feel if students use calculators too early, they won't learn math concepts

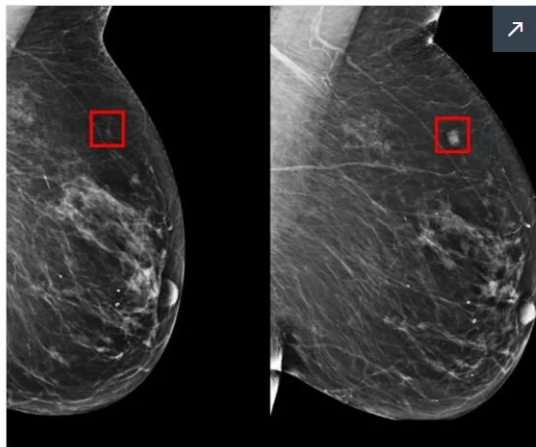
Math teachers protest against calculator use

By JILL LAWRENCE

"My older kids don't pay any attention to an answer being absurd," he said. "Teachers are shy."

Automatizar tareas Trabajo?





TECNOLOGÍA

La inteligencia artificial puede ayudar a un diagnóstico temprano de cáncer de mama

sábado, 8 de julio de 2023



GUARDA

<https://twitter.com/i/status/1702240836564427145>





ROSALÍA, QUEVEDO & ZAID - AHORA TE JODE (IA) [Lyric Video]

221 K visualizaciones • hace 1 mes



KBYK

@z.AI.d @z.AI.d.