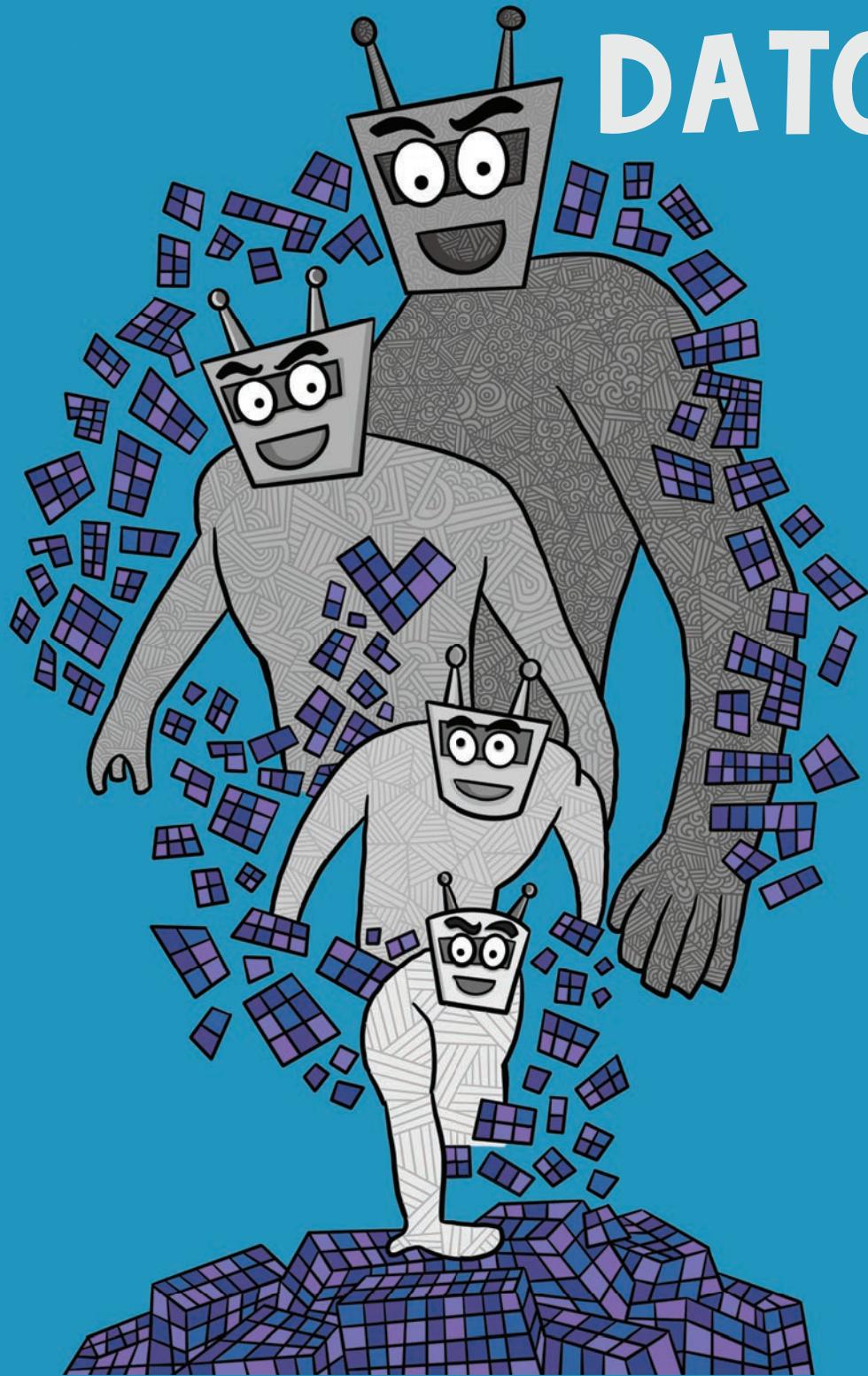


Somos IA n.º 2:

APRENDER DE LOS DATOS



© Julia Stoyanovich & Falaah Arif Khan (2022)

Traducido por Daniel Domínguez Figaredo

Términos de uso

Todos los contenidos gráficos/viñetas de este cómic están protegidos por una licencia CC BY-NC-ND 4.0. Consulte la página web de las licencias para obtener detalles sobre cómo puede usar este material gráfico.

Se puede usar paneles/grupos de paneles en presentaciones/artículos, siempre y cuando:

1. Se proporcione la cita adecuada.
2. No se realicen modificaciones a los paneles individuales.

Citar como:

Julia Stoyanovich y Falaah Arif Khan. "Aprender de los datos". We are AI Cómics, Vol. 2 (2021) <http://r-ai.co/comics>

CUANDO HABLAMOS DE INTELIGENCIA ARTIFICIAL (IA), GENERALMENTE HACEMOS UNA DISTINCIÓN ENTRE LOS SISTEMAS DE "IA CLÁSICA":

AQUELLOS QUE FUNCIONAN DE ACUERDO CON REGLAS PREDETERMINADAS, COMO ROOMBA, LA ASPIRADORA INTELIGENTE,



Y AQUELLOS QUE "APRENDEN" ESAS REGLAS A PARTIR DE LOS DATOS:

COMO LA IA PARA JUGAR AL AJEDREZ

O UN COCHE AUTÓNOMO.

PERO ESA DISTINCIÓN NO ES DEL TODO PRECISA.

DE HECHO, LA MAYORÍA DE LOS SISTEMAS DE IA QUE EXISTEN HOY EN DÍA COMBINAN REGLAS HECHAS A MANO CON ALGUNA FORMA DE AUTO-APRENDIZAJE.

VEAMOS ALGUNOS EJEMPLOS DE ESAS REGLAS Y CÓMO LAS MÁQUINAS LAS APRENDEN A PARTIR DE LOS DATOS.

SUPONGAMOS QUE TENEMOS QUE DISEÑAR UN SISTEMA DE ILUMINACIÓN INTELIGENTE



ESTE ALGORITMO ES MUY SIMPLE, Y TIENE UN SOLO PASO:

TOMA LAS CONDICIONES EXTERIORES (SI ESTÁ OSCURO AFUERA) COMO ENTRADA,

Y UTILIZA LA REGLA 1 PARA PREDICIR LA SALIDA ADECUADA: ENCENDER LA LUZ O NO.

ESTA REGLA ES SIMPLE, PERO ¿ES BUENA?

¿REPRESENTA ADECUADAMENTE LA FORMA EN QUE ENCENDEMOS MANUALMENTE LAS LUCES EN CASA?

¡HAGAMOS UN EXPERIMENTO PARA AVERIGUARLO!

RECOGEREMOS OBSERVACIONES SOBRE SI ESTÁ OSCURO AFUERA Y SI LAS LUCES ESTÁN ENCENDIDAS EN LA CASA.

A

PARA CADA OBSERVACIÓN, COMPROBAREMOS SI LA PREDICCIÓN REALIZADA POR NUESTRO ALGORITMO COINCIDE CON LO QUE OBSERVAMOS.

EN LAS OBSERVACIONES A Y B, ESTÁ OSCURO AFUERA, PERO ESTAMOS VIENDO LA TELEVISIÓN Y CENANDO,



ASÍ QUE LAS LUCES ESTÁN ENCENDIDAS.

¡Y ESO ES, DE HECHO, LO QUE HA PREDICHO LA REGLA 1!

ESTO SIGNIFICA QUE TANTO A COMO B RESPALDAN LA HIPÓTESIS CODIFICADA EN LA REGLA 1.

B

LA OBSERVACIÓN C TAMBIÉN APOYA LA HIPÓTESIS:

ESTÁ DESPEJADO Y SOLEADO, ASÍ QUE LAS LUCES ESTÁN APAGADAS.

C

VEAMOS AHORA LA OBSERVACIÓN D:

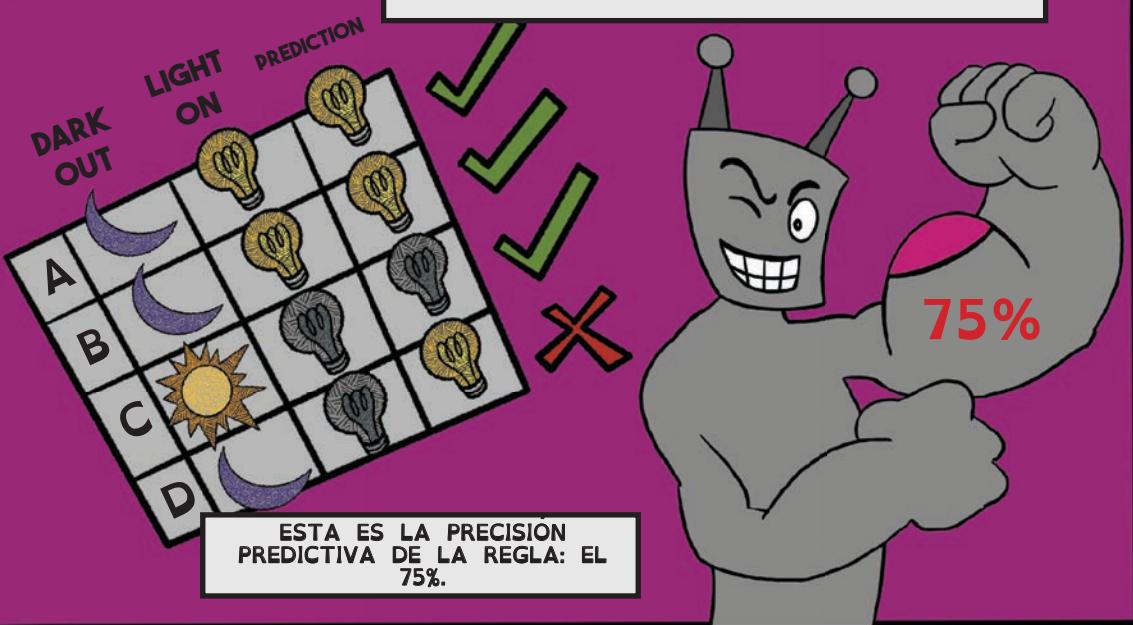
ESTÁ OSCURO AFUERA Y ES HORA DE DORMIR, ASÍ QUE LAS LUCES ESTÁN APAGADAS.

D

PERO, SEGÚN LA REGLA 1, EL ALGORITMO HA PREDICHO QUE LAS LUCES ESTARÍAN "ENCENDIDAS". ¡LA REGLA 1 FALLA AQUÍ!

VEAMOS LO QUE HEMOS APRENDIDO SOBRE LA REGLA 1 DE NUESTRO EXPERIMENTO:

NUESTRA REGLA PUDO PREDICIR CORRECTAMENTE LA SALIDA 3 DE 4 VECES.



HAGAMOS ALGUNAS ACLARACIONES SOBRE EN LA TERMINOLOGÍA EN ESTE CASO:

HABITUALMENTE, NOS REFERIMOS A LA ENTRADA DE LA REGLA (SI ESTÁ OSCURO, EN NUESTRO EJEMPLO) COMO LA CARACTÉRISTICA DE ENTRADA, O SIMPLEMENTE CARACTERÍSTICA.

EL ESTADO EN EL QUE SE ENCUENTRAN NUESTRAS LUces (ENCENDIDO/APAGADO) ES EL RESULTADO, Y GENERALMENTE SE DENOMINA RESULTADO O ETIQUETA.



LA REGLA CON LA QUE ESTAMOS TRABAJANDO SE LLAMA "CLASIFICADOR": ASIGNA UNA ETIQUETA DE CLASE A UNA OBSERVACIÓN.

EL CLASIFICADOR QUE DISEÑAMOS USANDO LA REGLA 1 PUEDE HACER UNA DE ESTAS DOS OPCIONES: ENCENDER O APAGAR LAS LUces.

TALES CLASIFICADORES SE LLAMAN "BINARIOS".

VOLVAMOS AHORA A LA OBSERVACIÓN D, QUE LA REGLA 1 CLASIFICÓ INCORRECTAMENTE:

¿CUÁL ES EL COSTE DE ESE ERROR?

PUES, QUE LA LUZ INTELIGENTE NOS DESPIERTE EN MEDIO DE LA NOCHE.

¿QUÉ PODEMOS HACER AL RESPECTO?

SI COMPARAMOS LAS OBSERVACIONES A, B, Y D,

VEMOS QUE TIENEN EL MISMO VALOR DE LA CARACTERÍSTICA DE ENTRADA (ESTÁ OSCURO FUERA), PERO SALIDAS DIFERENTES.

ESTO NOS HACE PENSAR QUE NECESITAMOS ALGUNAS CARACTERÍSTICAS ADICIONALES PARA DISTINGUIR ENTRE ESTAS SITUACIONES.

AHORA BIEN, COMO DISEÑADORES DE ESTE SISTEMA, DEBEMOS HACERNOS UNA PREGUNTA:

¿QUÉ OTRAS ENTRADAS PUEDEN SER ÚTILES?

SEGÚN TU EXPERIENCIA, ¿QUÉ CARACTERÍSTICAS CREEES QUE SERÁN ÚTILES PARA PREDICIR EL RESULTADO?

¿LA TEMPERATURA EXTERIOR? ¡NO!

¿EL PRECIO DEL TÉ EN CHINA? ¡NO!

QUIZÁS, ¿ES HORA DE ACOSTARSE?

PONGÁMONOS A TRABAJAR PARA REFINAR NUESTRA REGLA.

ESTA ES NUESTRA REGLA
REFINADA:

LA LLAMAREMOS
REGLA 2.



"ENCENDER LAS
LUCES SI ESTÁ OS-
CURO AFUERA,



DE LO CONTRARIO, MANTENER
LAS LUCES APAGADAS."

Y AÚN NO ES HORA
DE ACOSTARSE.

VOLVAMOS A EJECUTAR NUESTRO
EXPERIMENTO, RECOPELEMOS ALGUNAS
OBSERVACIONES MÁS Y EVALUEMOS LA
NUESTRA NUEVA REGLA.

E



¡YA ESTÁ! ¡PARECE QUE
FUNCIONA PERFECTAMENTE!

F

NO NECESITAMOS LAS LUCES ENCENDIDAS A
MENOS QUE ESTÉ OSCURO AFUERA Y ESTEMOS
DESPERIERTOS HACIENDO ALGO!

G



H



Y ESTA VEZ, INCLUSO, EL GENIO APAGA LAS LUCES SI NOS
QUEDAMOS DORMIDOS EN EL SOFÁ (DESPUÉS DE LA HORA DE
ACOSTARNOS), ¡QUÉ CARIÑOSO!

LA REGLA 2 HA DEMOSTRADO UNA PRECISIÓN PERFECTA:

NUESTRO ALGORITMO ES CAPAZ DE PREDICIR EL RESULTADO CORRECTAMENTE EL 100 % DE LAS VECES.



HEMOS UTILIZADO NUESTRA INTUICIÓN Y EXPERIENCIA PARA DISEÑAR UNA REGLA ADECUADA, DATOS RECOPILADOS PARA VALIDAR SU RENDIMIENTO,

E HICIMOS LOS AJUSTES NECESARIOS HASTA QUE ALCANZAMOS UN RENDIMIENTO CON EL QUE ESTAMOS SATISFECHOS.

DE HECHO, ESTAMOS TAN CONTENTOS CON NUESTRA LUZ INTELIGENTE,

CON LO BIEN QUE FUNCIONA Y CÓMO NOS AYUDA A CONSERVAR ENERGÍA,

QUE NOS GUSTRARÍA OFRECERLA A OTROS PARA QUE TAMBIÉN LA USEN.

VEAMOS SI FUNCIONA PARA QUE NUESTRA AMIGA ANN LA USE EN SU OFICINA.

PERO ANTES DE SEGUIR ADELANTE Y VOLVER A CABLEAR TODO EL EDIFICIO DE OFICINAS DE ANN, RECOPILEMOS ALGUNOS PATRONES DE USO PARA VER SI NUESTRA REGLA SERÍA EFECTIVA EN ESTE NUEVO CASO.

YA CONOCEMOS EL EJERCICIO: ¡HAGAMOS UN EXPERIMENTO!

RECORDEMOS LA REGLA 2:

"ENCENDER LAS LUCES SI ESTÁ OSCURO AFUERA Y TODAVÍA NO ES LA HORA DE ACOSTARSE. DE LO CONTRARIO, MANTENER LAS LUces APAGADAS".

I



EN LA OFICINA, LAS LUces ESTÁN ENCENDIDAS DURANTE EL DÍA CUANDO LAS PERSONAS TRABAJAN.

PERO LA REGLA 2 LAS APAGARÍA INCORRECTAMENTE, YA QUE HAY LUZ AFUERA.

J



LA MAYORÍA DE LAS TARDES, CUANDO YA TODO SE HAN IDO A CASA, LAS LUces SE APAGAN

K



PERO LA REGLA 2 LAS ENCENDERÍA INCORRECTAMENTE, YA QUE ESTÁ OSCURO!

L



LOS FINES DE SEMANA, LA OFICINA ESTÁ VACÍA Y LAS LUces ESTÁN APAGADAS DURANTE TODO EL DÍA.

EN ESTE CASO, LA REGLA 2 FUNCIONA CORRECTAMENTE: MANTIENE LAS LUces APAGADAS, YA QUE ES DE DÍA AFUERA.

CADA CIERTO TIEMPO, ALGUNI SE QUEDA HASTA TARDE EN LA OFICINA, TERMINADO LAS TAREAS PARA UNA FECHA DE ENTREGA IMPORTANTE.

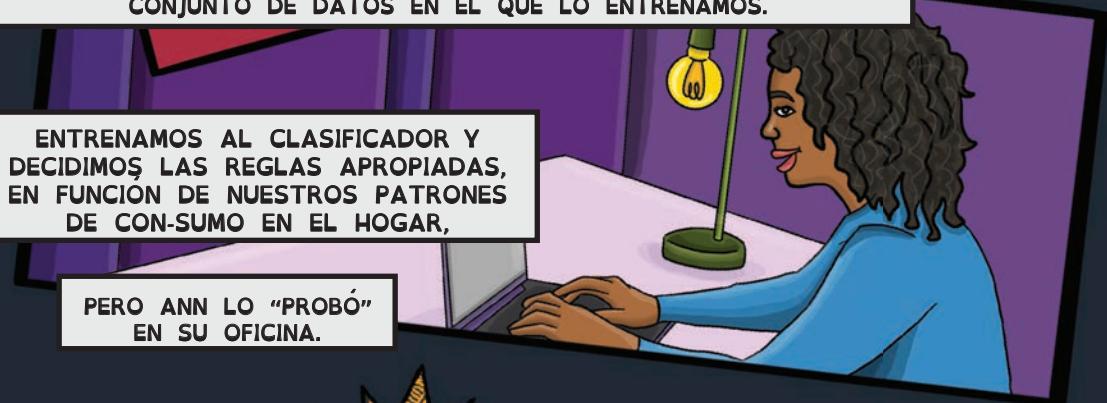
NECESITAMOS LAS LUces ENCENDIDAS EN ESTA SITUACIÓN, PERO LA REGLA 2 FALLA DE NUEVO: APAGARÍA LAS LUces, YA QUE ES TARDE (Y YA HA PASADO LA HORA DE ACOSTARSE).

¡OH, NO! ESO SIGNIFICA QUE SOLO OBTUVIMOS 1 PREDICCIÓN CORRECTA DE 4.

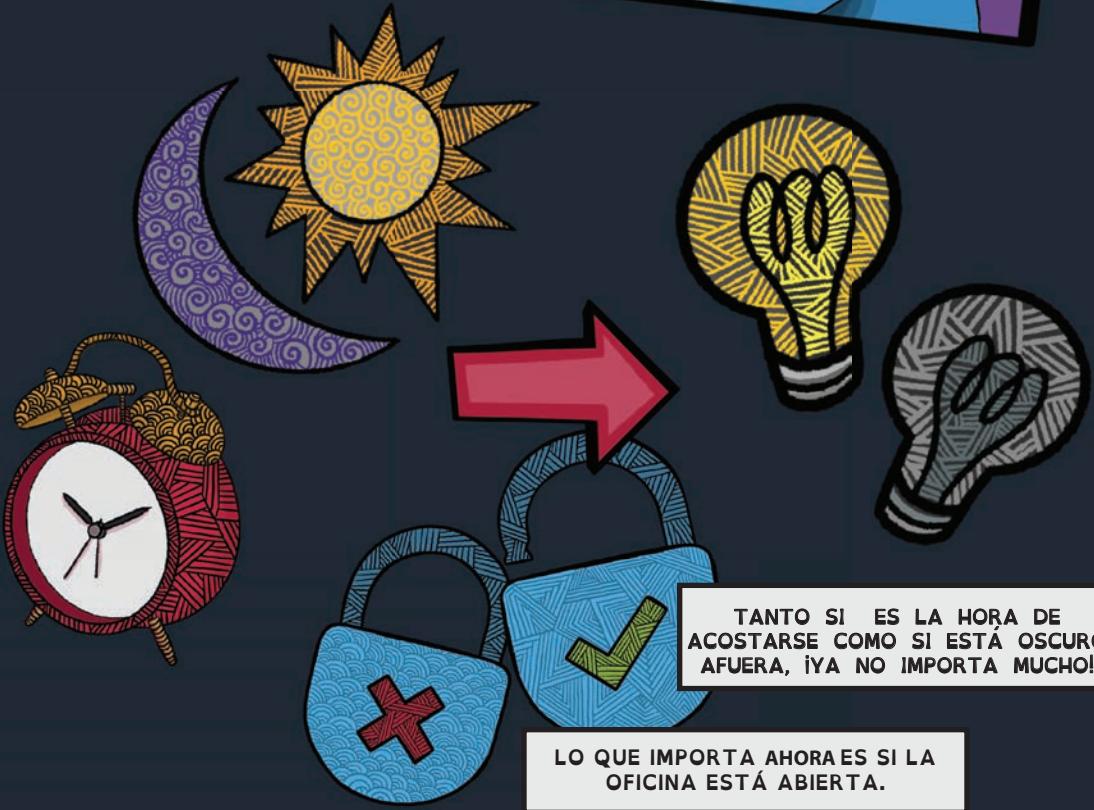
SI SUPONEMOS QUE LAS LUCES ESTÁN ENCENDIDAS APROXIMADAMENTE LA MITAD DEL TIEMPO, BASTA CON LANZAR UNA MONEDA AL AIRE PARA DECIDIR SI ENCENDER LAS LUCES O NO.



PORQUE EL CONJUNTO DE DATOS EN EL QUE PROBAMOS EL CLASIFICADOR REPRESENTA UNA SITUACIÓN DIFERENTE A LA DEL CONJUNTO DE DATOS EN EL QUE LO ENTRENAMOS.



PERO ANN LO "PROBÓ"
EN SU OFICINA.



¿CUÁL ES EL COSTO DE UN ERROR AQUÍ?

EL COSTO DE UN FALSO POSITIVO (ENCENDER LA LUZ INNECESSARIAMENTE CUANDO DEBERÍA ESTAR APAGADA) ES QUE DESPERDICIAMOS ENERGÍA Y PROVOCA UNA FACTURA DE ELECTRICIDAD CONSIDERABLE PARA LA COMPAÑÍA.



EL COSTE DE UN FALSO NEGATIVO (APAGAR LA LUZ CUANDO SE NECESITA) ES QUE INTERRUMPIMOS A LAS PERSONAS EN MEDIO DE SU TRABAJO Y SU PRODUCTIVIDAD SE VE AFECTADA.

¿SÉ QUÉ DEBEMOS HACER? VOLVER AL TABLERO DE DIBUJO OTRA VEZ?

¿PENSAR EN QUÉ CARACTERÍSTICAS ADICIONALES USAR, RECOPILAR ESEOS DATOS,

Y REALIZAR UN EXPERIMENTO PARA VERIFICAR SI NUESTRA REGLA REALMENTE FUNCIONA?

PERO ESO ES MUY TEDIOSO.

Y LO QUE ES MÁS IMPORTANTE, NOS RESULTARÁ MUY DIFÍCIL CONTINUAR AJUSTANDO ESTAS REGLAS QUE SON CADA VEZ MÁS COMPLEJAS,

SEGÚN LAS UBICACIONES DONDE SE UTILICE LA LUZ INTELIGENTE Y SUS DIFERENTES REQUISITOS.

LA RESPUESTA CORRECTA ES: USAR UN ALGORITMO DE APRENDIZAJE AUTOMÁTICO PARA IDENTIFICAR PATRONES ESTADÍSTICOS EN NUESTROS DATOS, Y "APRENDER" LAS REGLAS DE FORMA AUTOMÁTICA.

AL IGUAL QUE DISEÑAMOS EXPERIMENTOS PARA VERIFICAR SI LAS REGLAS QUE HABÍAMOS CREADO ERAN BUENAS,

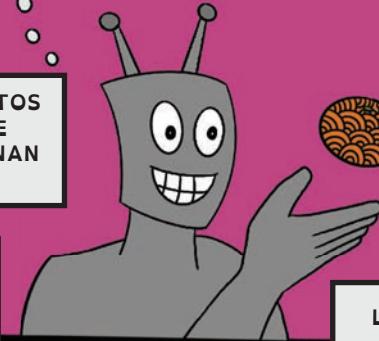


TAMBIÉN DEBEMOS DISEÑAR EXPERIMENTOS PARA PROBAR SI LAS REGLAS QUE SE APRENDEN AUTOMÁTICAMENTE FUNCIONAN BIEN CUANDO SE IMPLEMENTAN.

SI ESOS DATOS SON REPRESENTATIVOS DEL USO FUTURO DEL SISTEMA DE LUZ INTELIGENTE,

ENTONCES LA REGLA FUNCIONARÁ BIEN.

ESTO SE DEBE A QUE LA REGLA APRENDIDA DEPENDE COMPLETAMENTE DE LOS DATOS HISTÓRICOS:



PERO SI EL CASO DE USO REAL ES DIFERENTE,

LA REGLA COMETERÁ MUCHOS ERRORES,



COMO SUCEDE AL ENCENDER LA LUZ EN EDIFICIOS DE OFICINAS VACÍOS,

O AL DESPERTAR A UN NIÑO PEQUEÑO QUE DUEME .

TANTO SI LAS REGLAS SON ESCRITAS POR HUMANOS COMO SI SON APRENDIDAS POR LAS MAQUINAS A PARTIR DE LOS DATOS HISTÓRICOS, DEBEMOS ASEGURARNOS DE USAR ADECUATAMENTE EL MÉTODO CIENTÍFICO:

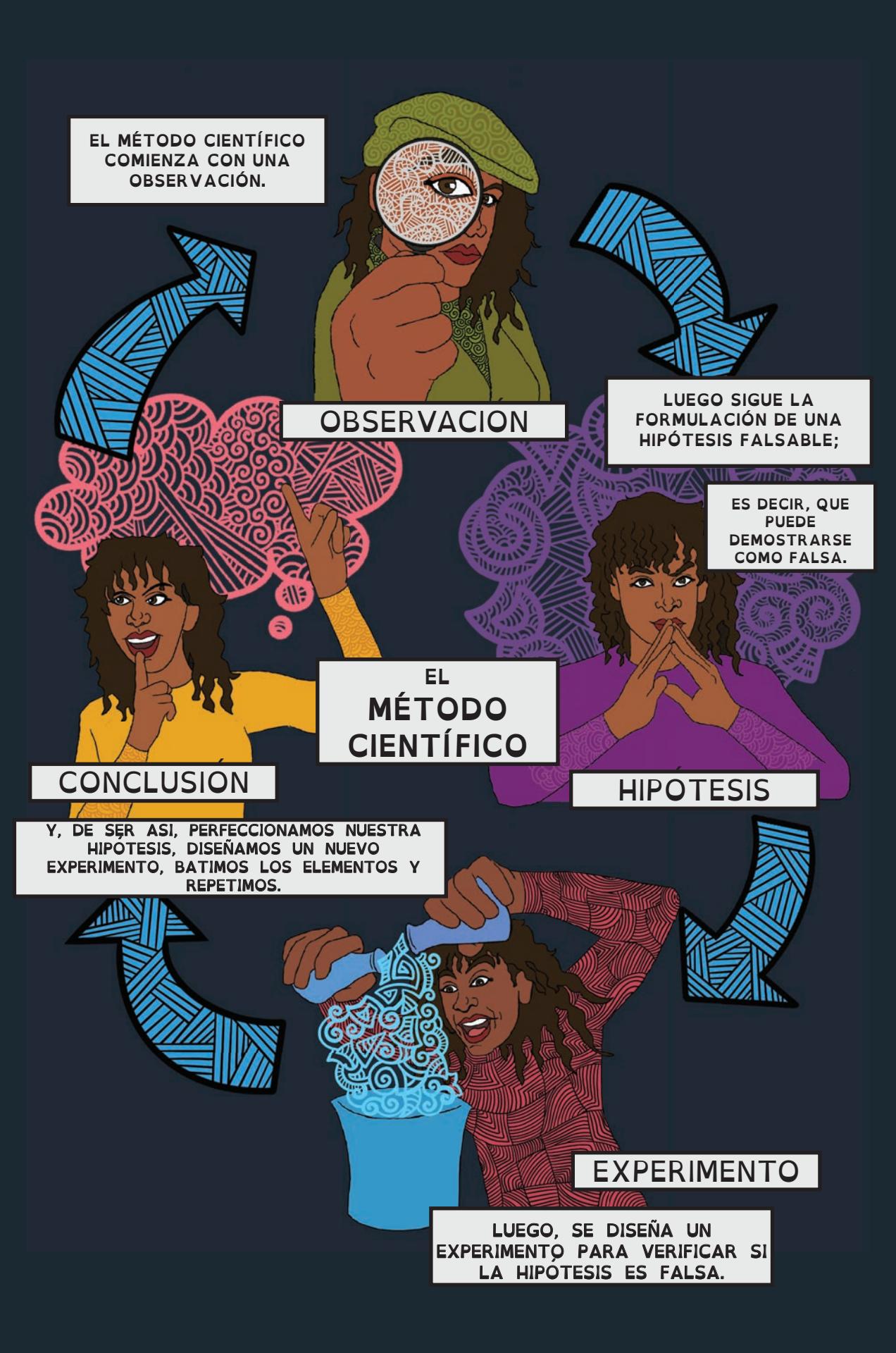


LA FORMULACIÓN Y LA PRUEBA DE UNA HIPÓTESIS FALSABLE.



CÓMO DIJO EL FILÓSOFO DE LA CIENCIA KARL POPPER:

"NINGUNA TEORÍA O IDEA DEBERÍA CONSIDERARSE CIENTÍFICA A MENOS QUE, EN PRINCIPIO, PUEDA DEMOSTRARSE SU FALSEDAD".



EL MÉTODO CIENTÍFICO COMIENZA CON UNA OBSERVACIÓN.



LUEGO SIGUE LA FORMULACIÓN DE UNA HIPÓTESIS FALSABLE;

ES DECIR, QUE PUEDE DEMOSTRARSE COMO FALSA.

EL MÉTODO CIENTÍFICO

CONCLUSIÓN

HIPOTESIS

Y, DE SER ASÍ, PERFECCIONAMOS NUESTRA HIPÓTESIS, DISEÑAMOS UN NUEVO EXPERIMENTO, BATIMOS LOS ELEMENTOS Y REPETIMOS.



LUEGO, SE DISEÑA UN EXPERIMENTO PARA VERIFICAR SI LA HIPOTESIS ES FALSA.

LA PREGUNTA BÁSICA: "¿FUNCIONA EL CLASIFICADOR?", SE FORMULA COMO "¿LAS PREDICCIÓNES DEL CLASIFICADOR SON MÁS PRECISAS DE LO QUE SERÍAN UNAS PROPUESTAS ALEATORIAS?"

ESTE ES EL NIVEL MÁS BAJO PARA CONSIDERAR LA PRECISIÓN:



NO QUEREMOS CONSTRUIR UNA IA SOFISTICADA PARA TERMINAR LANZANDO MONEDAS AL AIRE.

Y SI LA HIPÓTESIS RESULTA SER FALSA, NO LA USAMOS.

CUANDO TENEMOS INFORMACIÓN VERDADERAMENTE FIABLE SOBRE EL PAPEL (OBSERVACIONES DE SI LA LUZ DEBERÍA O NO ESTAR ENCENDIDA), ENTONCES PODEMOS VERIFICAR.

DISEÑAMOS UN NUEVO PROCEDIMIENTO PARA GENERAR DE DATOS Y VOLVEMOS A ENTRENAR EL MODELO PARA APRENDER UNA NUEVA REGLA.



ES IMPORTANTE DESTACAR QUE, INCLUSO CON TODA ESTA SOFISTICACIÓN ADICIONAL, ES POCO PROBABLE QUE ENCONTREMOS LA PRECISIÓN PERFECTA.



UNA DE LAS RAZONES POR LAS QUE ESTO OCURRE ES QUE EL MUNDO ESTÁ LLENO DE INCERTIDUMBRES.

OTRA ES QUE A VECES LAS REGLAS SE ROMPEN: LAS PERSONAS PUEDEN VENIR A TRABAJAR DURANTE EL FIN DE SEMANA O PERMANECER EN EL TRABAJO MÁS TIEMPO PARA CUMPLIR CON UN PLAZO DE ENTREGA.

EN SITUACIONES DE LA VIDA REAL, LO MEJOR QUE PODEMOS ESPERAR, ES QUE EL CLASIFICADOR FUNCIONE LA MAYOR PARTE DEL TIEMPO.



PERO DE VEZ EN CUANDO, ALGUIEN DEBERÁ LEVANTARSE Y ENCENDER O APAGAR LA LUZ.