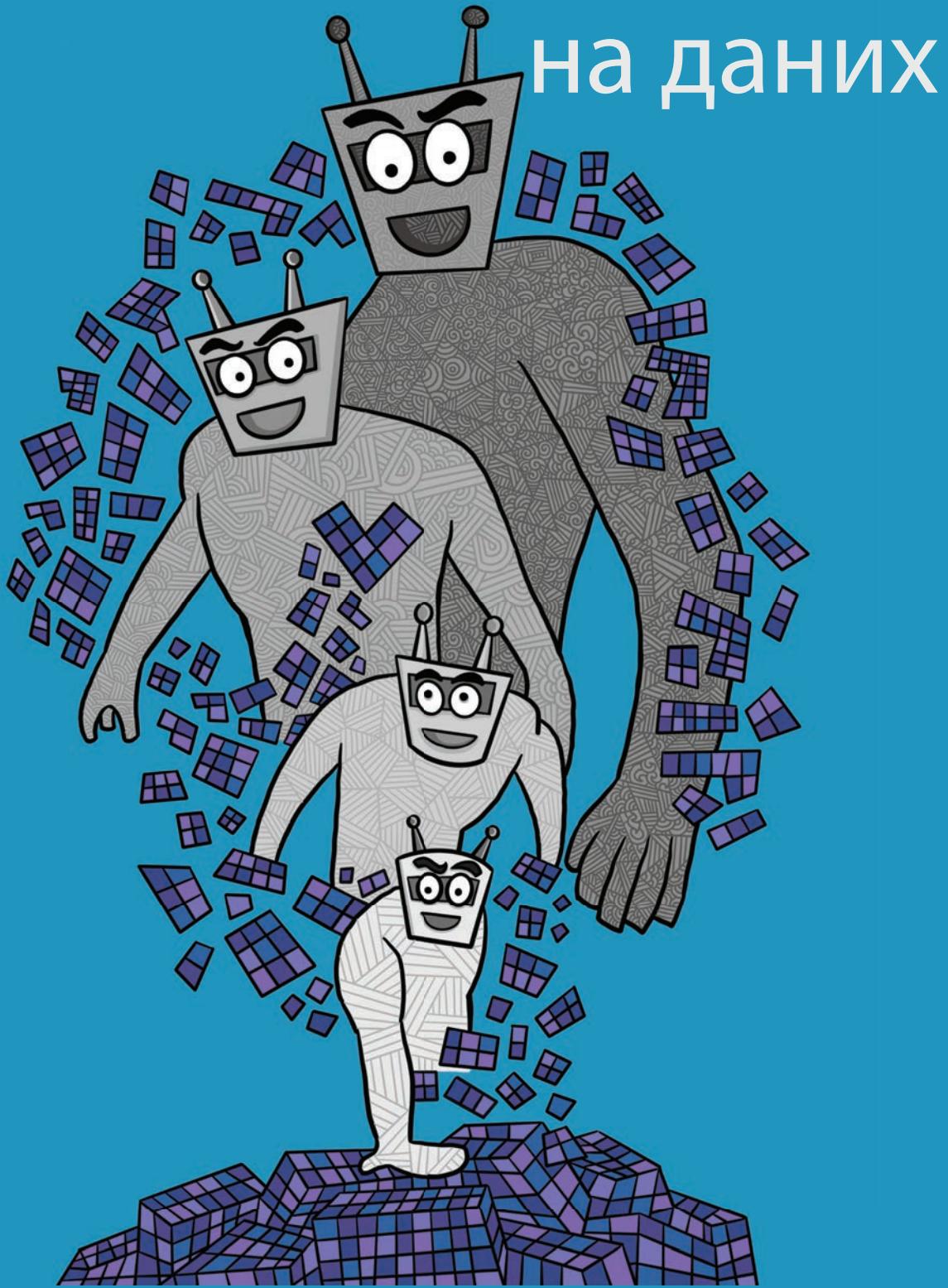


Ми і є ШІ № 2: Навчаючись на даних



© Julia Stoyanovich & Falaah Arif Khan (2022)
Translated by Yaroslava Savosh-Davydova; edited by Yevhen
Redko and Tetiana Zakharchenko.

УМОВИ ВИКОРИСТАННЯ

Усі ілюстрації в цьому коміксі доступні за ліцензією CC BYNC-ND 4.0. Будь ласка, перейдіть на сторінку ліцензії, щоб дізнатися більше про те, як можете використовувати ці роботи.

TL; DR: Не соромтеся використовувати панелі/групи панелей у презентаціях/статтях, якщо

1. належно цитуєте їх;
2. не вносите змін в окремі панелі.

Цитувати за:

Джулія Стоянович та Фала Аріф Хан. «Навчаючись на даних?». Ми і є ШІ. Комікси, том 2 (2021)

https://dataresponsably.github.io/we-are-ai/comics/vol2_en.pdf.

Контакти:

Будь ласка, надсилайте будь-які запити щодо використання елементів цього коміксу на адресу themachinelearnist@gmail.com та копію на stoyanovich@nyu.edu.

Говорячи про штучний інтелект (ШІ), ми зазвичай розрізняємо системи «класичного ШІ» —

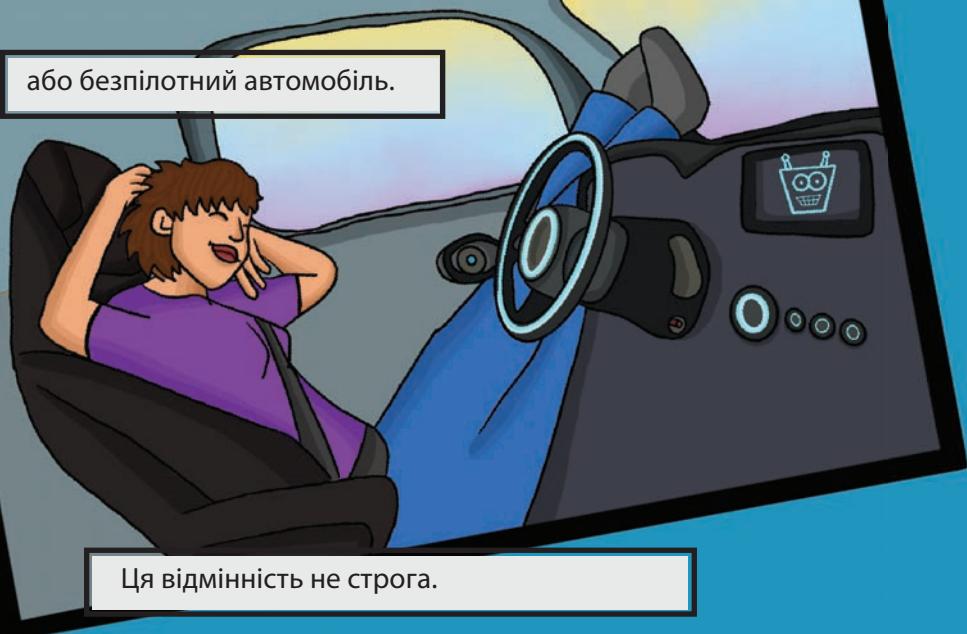
ті, які працюють за заздалегідь визначеними правилами, як от Roomba — розумний пилосмок,



і ті, які «вивчають» правила з даних, —

наприклад, штучний інтелект, що грає в шахи,

або безпілотний автомобіль.



Ця відмінність не строга.

Насправді більшість сучасних систем штучного інтелекту поєднують правила, створені вручну, з певною формою навчання.

Розгляньмо приклади таких правил і подивімось, як машини вивчають їх на основі даних!

Уявіть, що маєте спроектувати розумну систему освітлення,

Як ваш ШІ має вирішувати, коли так діяти?

Ми можемо почати проєктувати таку систему, спираючись на власний повсякденний «інтелект».

Назвімо це Правилом 1. Воно просте:

яка автономно вмикає та вимикає світло у вашій оселі.

Вмикайте світло, коли надворі темно.

В іншому разі тримайте вимикач вимкненим.

Цей алгоритм дуже простий, він має лише один крок —

приймає зовнішні умови (чи темно надворі) як вхідні дані,

і, використовуючи Правило 1, прогнозує прийнятний результат — чи вмикати світло.

Це правило просте, але чи добре воно працює?

Чи правильно воно відображає те, як вручну керуєте світлом у вашому домі?

Виконаймо експеримент, щоб з'ясувати це!

Ми збиратимемо дані про те, чи темно надворі й чи ввімкнене світло в будинку.

A

Для кожного спостереження перевіримо, чи збігається прогноз за нашим алгоритмом із тим, що спостерігали.

У спостереженях а і b надворі темно, але ми дивимося телевізор і вечеряємо,



тому залишили світло.

I це справді те, що передбачало Правило 1!

Це означає, що а і b підтверджують гіпотезу, закодовану у Правилі 1.

B

Спостереження с теж підтверджує гіпотезу:

надворі любо і ясно, тож ми не вмикали світло.

C

Тепер погляньмо на заувагу d:

надворі темно, ми вкладаємося спати, світло вимкнули.

D

Однак, згідно з правилом 1, алгоритм передбачив, що світло буде «увімкнене». Правило 1 тут не працює!

Отже, що ми дізналися про Правило 1 з нашого експерименту —

Наше правило змогло правильно передбачити результат 3 рази з 4.



Підправмо тут ще трохи термінології.

Зазвичай вхідні дані правила — у нашому прикладі пітьма — мають назву вхідна ознака, або просто ознака.

Вихідні дані — стан наших світильників (увімкнено/вимкнено) — зазвичай називають результатом, або **міткою**.



Правило, з яким працюємо, називають «**класифікатор**» — воно надає спостереженню мітку класу.

Класифікатор, який ми створили за допомогою Правила 1, може зробити один із двох виборів — увімкнути або вимкнути світло.

Ми називаємо такі класифікатори «**бінарними**».

Повернімось до спостереження d,
яке хибно класифікувало правило 1.

Яка ціна цієї помилки?

Як ви, напевно, здогадалися,
розумне світло розбудило вас
посеред ночі!

Як можемо
зарадити?

Якщо порівнямо
спостереження a, b та d,

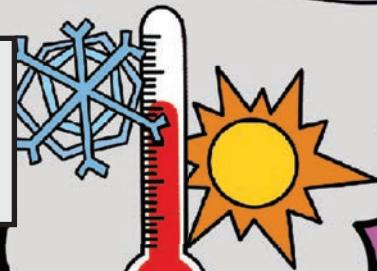
то побачимо, що вони мають однакове значення вхідної ознаки
(на вулиці темно), але різні вихідні дані.

Це наштовхує на думку, що потрібні додаткові ознаки, щоб
розділити такі ситуації.

Тепер рішення, яке ви, як розробник цієї
системи, маєте прийняти, стосується того,

які ще вхідні дані можуть
бути корисні?

Які ознаки, на вашу
думку, допоможуть
передбачити
результат?



Температура на
вулиці? Ні!

Ціна чаю в Китаї? Ні!

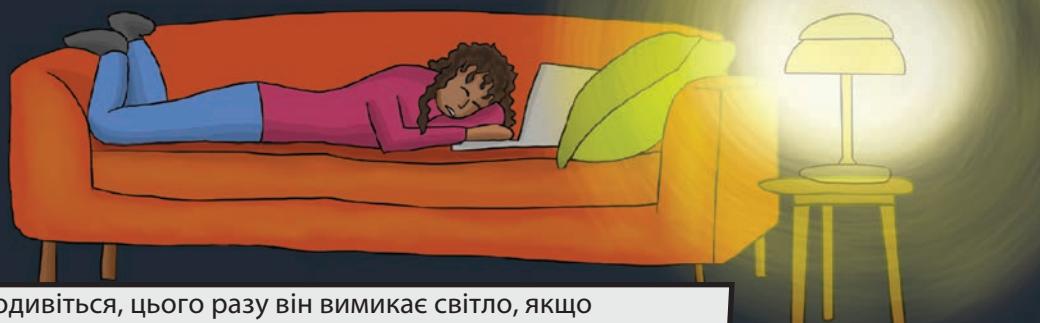
Гаразд, а як щодо
того, чи час спати?



Почнімо вдосконалювати наше
правило.

Ось наше вдосконалене
правило —

назвімо його
Правило 2.



І подивіться, цього разу він вимикає світло, якщо засинаємо на дивані (після того, як нам час спати) — так дбайливо!

Правило 2 продемонструвало бездоганну точність —

наш алгоритм здатен правильно передбачити результат у 100 % випадків!



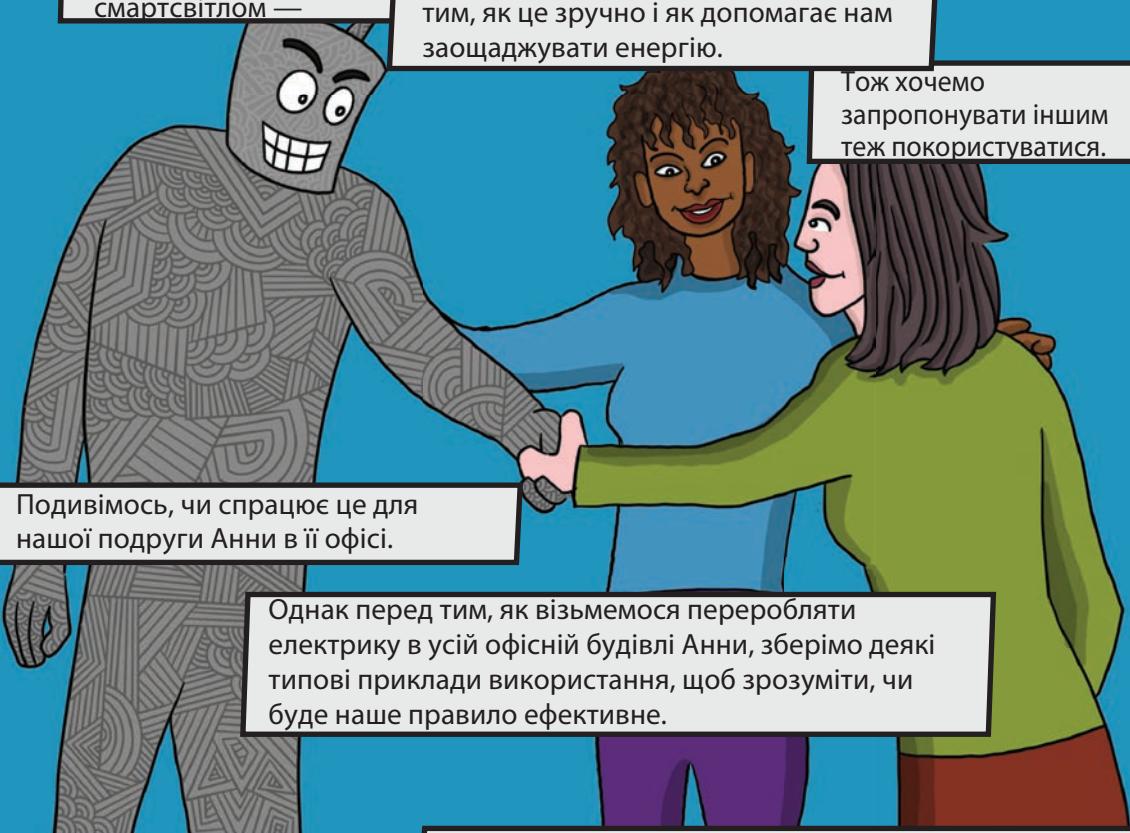
Ми використали інтуїцю й досвід, щоб розробити належне правило, а також зібрали дані, прагнучи перевірити, як воно спрацьовує.

і вносили потрібні корективи, аж доки досягли результату, який нам сподобався!

Насправді ми так задоволені нашим смартсвітлом —

тим, як це зручно і як допомагає нам заощаджувати енергію.

Тож хочемо запропонувати іншим теж покористуватися.



Ви знаєте правила: гайда експериментувати!

Пригадайте правило 2 —

«Вмикайте світло, якщо надворі темно, але ще не час спати. В іншому разі не вмикайте світла

I



В офісі ми тримаємо світло ввімкненим упродовж дня, коли люди працюють.

Однак Правило 2 некоректно спрацює, адже надворі світло

J



Переважно вечорами, коли всі розходяться по домівках, світло вимикається.

K



Однак Правило 2 некоректно ввімкне його, щойно на вулиці потемніє!

L

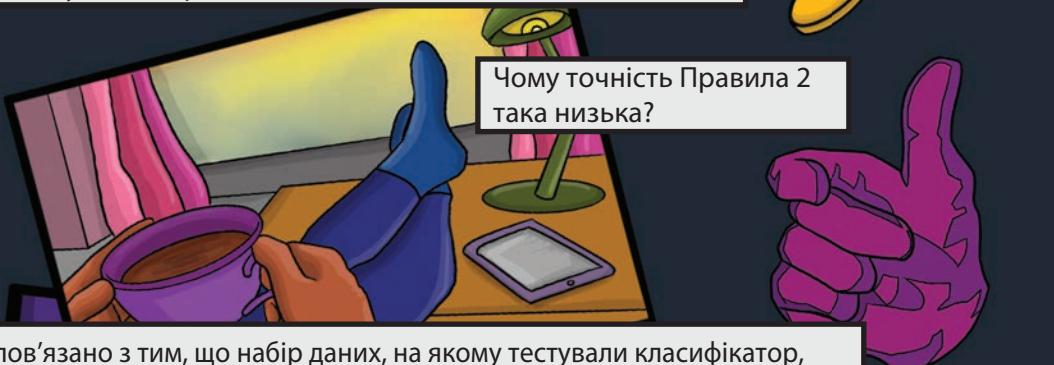
На вихідних офіс пустує, тому світло вимкнене протягом усього дня!

Час від часу хтось затримується в офісі допізна, щоб встигнути завершити роботу до критично важливого речення.

У цьому сценарії нам потрібно ввімкнути світло, але Правило 2 знову хибне: воно вимкне світло, адже вже пізно (після того, як зазвичай лягаємо спати).

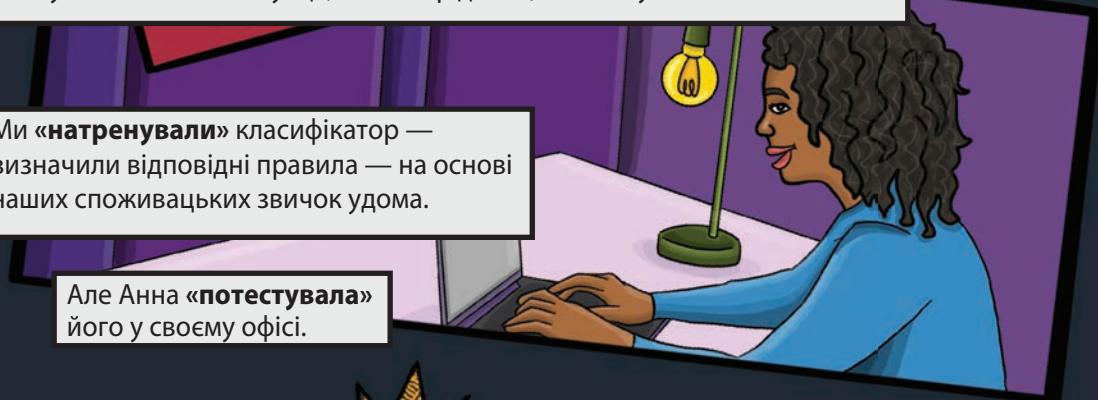
О, ні! Це означає, що ми отримали лише 1 правильний прогноз з 4!

Якщо припустити, що світло ввімкнене приблизно половину часу, то було б краще, якби ми просто підкинули монетку, щоб вирішити, чи вмикати світло!



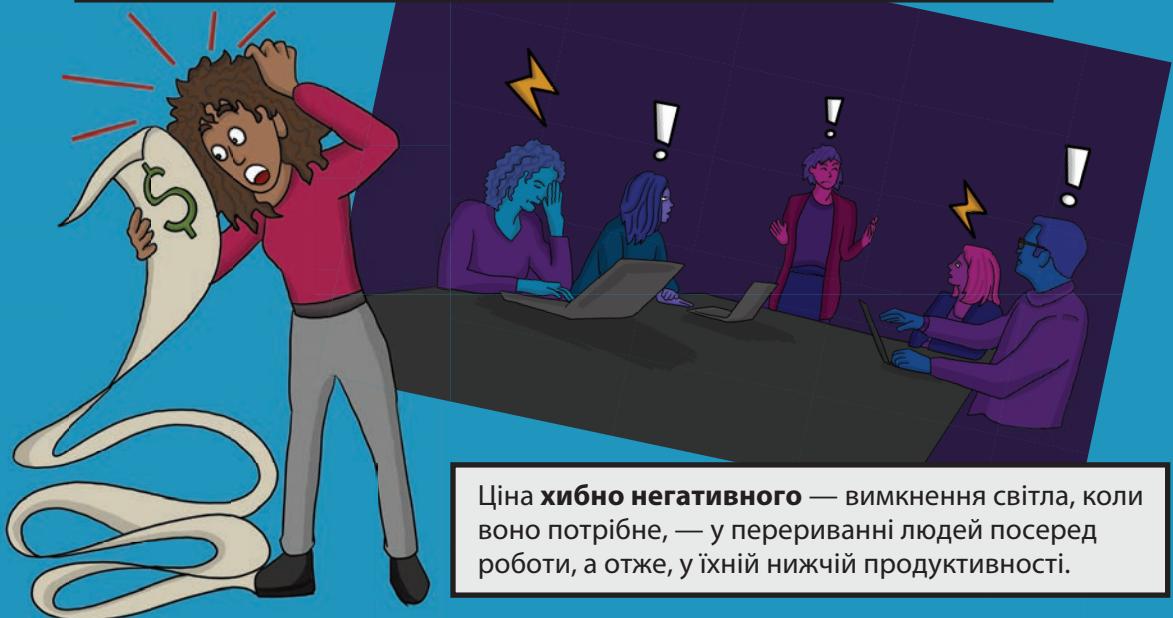
Це пов'язано з тим, що набір даних, на якому тестували класифікатор, стосується інакшої ситуації, ніж набір даних, на якому його навчали.

Ми «**натренували**» класифікатор — визначили відповідні правила — на основі наших споживацьких звичок у дома.



Яка тут ціна помилки?

Ціна хибно позитивного — невиправдане ввімкнення світла, коли воно має бути вимкнене, — змарнована електроенергія і кругленький рахунок компанії за неї.



Ціна хибно негативного — вимкнення світла, коли воно потрібне, — у перериванні людей посеред роботи, а отже, у їхній нижчій продуктивності.

Що нам робити? Знову повернутися на початок?

Подумайте, які додаткові ознаки потрібні, зберіть ці дані

і виконайте експеримент, щоб перевірити, чи працює наше правило насправді?

Але це так нудно!

І найголовніше — нам буде дуже складно далі формулювати ці правила, які стають щораз складнішими

залежно від місця, де використовують смартсвітло, та їхніх різних вимог!

Рішення таке: використайте алгоритм машинного навчання, щоб визначити статистичні закономірності в наших даних — і автоматично «навчіть» правил

Однак так само, як ми розробляли експерименти, щоб перевірити, чи застосовні вигадані правила на практиці,

важко розробити експерименти, щоб перевірити, чи працюють автоматично "навчені" правила в разі впровадження.

Бо правило повністю залежить від історичних даних —

якщо ці дані репрезентативні для майбутнього використання системи смартосвітлення,

тоді правило працюватиме добре.

Однак якщо фактичний випадок використання відрізняється —

правило призведе до багатьох помилок

Наприклад, увімкне світло в порожніх офісних будівлях,

або пробудить малюка серед денного сну.

Незалежно від того, правила записані людьми чи вивчені на основі історичних даних машинами, ми повинні обов'язково використовувати науковий метод —

формулювання та перевіряння спростованої гіпотези,

«Теорія або ідея не має бути науковою, якщо її не можна довести як хибну».

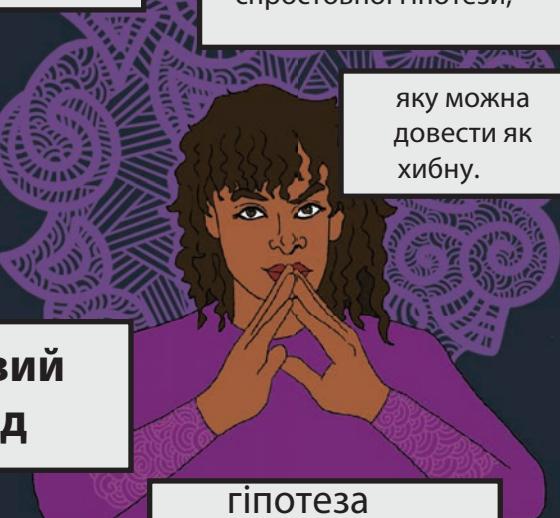
Як казав філософ науки Карл Поппер:



Науковий метод починається зі спостереження.

спостереження

Потім іде формулювання спростовної гіпотези,



яку можна довести як хибну.

науковий
метод

ВИСНОВОК

гіпотеза



І якщо так, ми уточнюємо гіпотезу, розробляємо новий експеримент, «прополіскуємо» його й повторюємо.

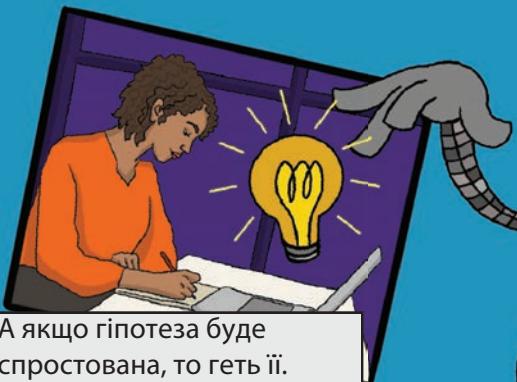


експеримент

Далі потрібно розробити експеримент, щоб перевірити, чи можна спростувати гіпотезу.

Ключове питання: «Чи працює класифікатор?» формалізоване як «Чи передбачення класифікатора точніші, ніж випадкове вгадування?»

Це найнижча планка для точності —



А якщо гіпотеза буде спростована, то геть її.

ми не хочемо зрештою створити дурненький ШІ, який підкидатиме монети.

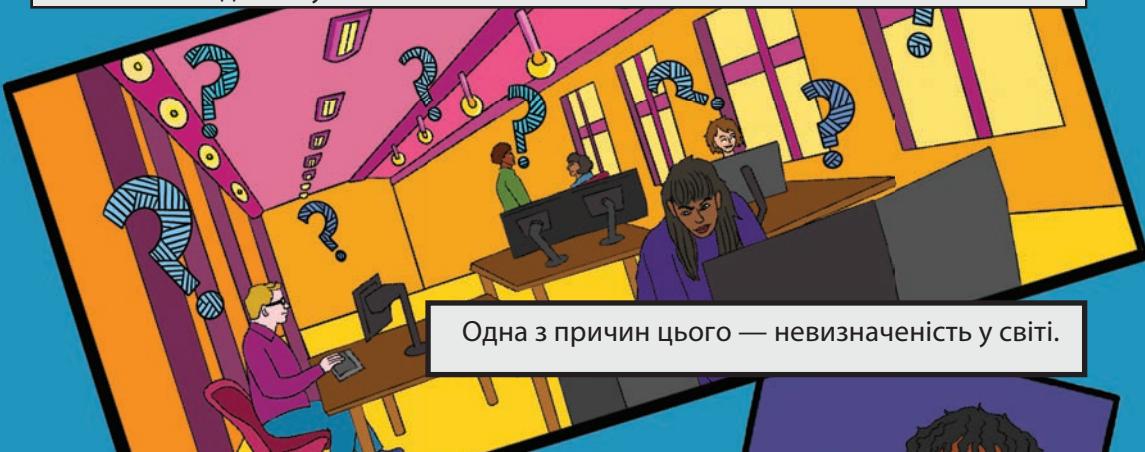


Коли маємо достовірну інформацію (спостереження за тим, чи справді має бути світло) — можемо перевірити.

Ми розробляємо нову процедуру генерування даних і перенавчаємо модель, щоб вона вивчила нове правило.



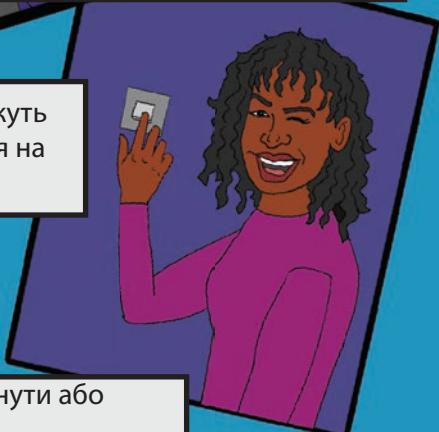
Важливо, що навіть з усією додатковою складністю ми однаково навряд чи побачимо ідеальну точність.



Одна з причин цього — невизначеність у світі.

А ще те, що правила іноді порушують: люди можуть виходити на роботу у вихідні або затримуватися на роботі довше, щоб вкластися в реченець.

Отже, у реальних життєвих ситуаціях можемо сподіватися лише на те, що класифікатор працюватиме більшість часу.



Але інколи хтось матиме встати й увімкнути або вимкнути світло.