Intel·ligència Artificial, mite o realitat?

vista prèvia >

La Intel·ligència Artificial ha passat en poques dècades de ser imaginable únicament en pel·lícules de ciència-ficció a formar part habitual del nostre dia a dia. Les apps dels telèfons mòbils, el reconeixement facial, les cerques a internet o els sistemes intel·ligents de gestió de residus formen part d'aquesta nova realitat tecnològica. Comprendre què significa la IA, saber com s'ha desenvolupat i com funciona, és un camí necessari per poder afrontar els reptes que, com tota revolució tecnològica, planteja de cara al futur.

Karina Gibert
Catedràtica de la Universitat Politècnica de Catalunya
Sotsdirectora del centre de recerca Intel·ligent Data Science and Artificial Intelligence (IDEAI-UPC)
@KarinagibertK

La popular IA, els orígens

En els últims temps no fem més que sentir parlar d'Intel·ligència Artificial (IA), de *big data*, d'internet de les Coses, de *blockchain*, i de tota mena d'aplicacions més que a la immensa majoria de la ciutadania pot ser que els hi soni més com a ciència-ficció que una altra cosa. Llavors... Què és la IA? És tan mítica com sembla? Tan poderosa o invasiva com la pinten? Ens dominarà? Ens millorarà les vides, les feines i els espais on existim?

El cert és que la IA està de moda, i genera actualment un aura de misteri, intriga i curiositat en una gran part de la població que en sent parlar per tot arreu a totes hores sense saber gaire de què es tracta. L'objectiu d'aquest article és posar una mica de llum i situar què és la IA més enllà d'imatges preconcebudes de robots i androides. De fet, estem davant d'una de les poquíssimes disciplines científiques que té un origen definit, situat, conegut i perfectament determinat. Al 1956, el científic nord-americà John McCarthy (1927-2011) convoca una escola d'estiu a Darmouth, als EUA, on ell era professor, a la que hi assisteix la flor i nata del moment: des d'acadèmics i investigadors com Marvin L. Minsky (1927-2016) de la Harvard University, Herbert Simon (1916-2001) i Allen Newell (1927-1992) del Carnegie-Mellon Collage o Oliver Selfridge (1926-2008), Ray Solomonoff (1926-2009), i fins a enginyers i directius d'empreses com Nathanel Rochester (1919-2001), Arthur Samuel (1901-1990) o Trenchard More (1930-2019) d'IBM, Claude Shannon (1916-2001) de Bell... En total deu eminències concentrades per dos mesos en un estudi al voltant d'una conjectura que assegurava que «les activitats intel·ligents es poden descriure amb tanta precisió que es pot fer una màquina que les simuli».[1] I així neix la IA, com «una ciència empírica que estudia els mecanismes de la intel·ligència humana», segons el propi McCarthy.

De símbols i subsímbols: El gran cisma

La IA té un origen bicèfal, i es debat entre dos grans enfocaments que determinen la naturalesa de tota la recerca que s'ha fet fins a l'actualitat. D'una banda, hi ha el que anomenem la Intel·ligència Artificial simbòlica, encapçalada pel propi McCarthy, que posa el focus en aconseguir que les màquines imitin la intel·ligència humana. És a dir, que puguin fer coses per les quals es requereix intel·ligència, però no només que les facin, sinó que les portin a terme «com les faria un humà». I així, en aquesta branca de la disciplina, el principal interès se situa en el procés de com es resolen els problemes, o com es desenvolupa l'acció, i en construir models de raonament automàtic que mimetitzin, salvant les distàncies, el dels humans. Aquesta és una branca de la IA on la lògica formal té molt pes i on s'intenta «entendre el pensament» humà i imitar-lo.

Per l'altra banda, la IA connexionista —o subsimbòlica— hereva directa de la cibernètica, i encapçalada per Minsky, defineix la IA com «la ciència de fer màquines que facin el que fan els humans com a éssers intel·ligents».[2] En aquest cas, el focus es posa més aviat en dissenyar algo-

1 MCCARTHY (et al.), «A proposal for the dartmouth summer research project on artificial intelligence».

2 MINSKY, *Heuristic aspects of the artificial intelligence problem*.

A la IA connexionista, el focus es posa més aviat en trobar solucions semblants a les que obté un humà encara que la forma com s'arribi a la solució sigui diferent a la de l'humà

ritmes[3] que donin resultats semblants als que obté un humà quan fa i desfà, però on la forma com es construeix aquesta solució «compatible» amb la que donaria l'humà ja no és tan important. Per entendre'ns, des d'aquesta perspectiva, l'important era que Deep Blue —el supercomputador d'IBM— guanyés a Garry Kasparov (1963) als escacs el 1996, i no tant si l'andròmina raonava com un jugador d'escacs humà per decidir quina peça bellugar en el següent moviment i cap a on; o explorant espais combinatoris de solucions possibles amb milers de càlculs més o menys ordenats que res tenen a veure amb com juguen els professionals dels escacs. Des d'aquesta branca, la IA es veu més aviat com una peça per augmentar el pensament humà i no es posa tant d'interès —més aviat gens— per comprendre'l. I així, la IA subsimbòlica genera algoritmes amb molts càlculs, que anomenem «de caixa negra», perquè donen solucions que no es poden justificar de cap forma comprensible per l'ésser humà. De fet, molts mètodes d'aquesta branca, com les xarxes neuronals artificials,[4] es basen en metàfores biològiques del més baix nivell d'intel·ligència, que connecten càlculs segons les regles del sistema biològic de referència, a l'espera que el comportament intel·ligent es produeixi de forma espontània amb les interaccions dels diferents components. De forma similar a com cada formiga fa una petita tasca dins el funcionament del formiguer.

El curiós és que aquestes dues visions s'han mantingut des dels orígens fins avui dia, a l'extrem que recentment ha emergit una nova branca de la IA que s'anomena «*explainable AI*»,[5] que intenta precisament assegurar que totes les IA podran explicar el perquè de les solucions que proposen, propietat que ja tenia la IA simbòlica i que encara avui representa un repte descoratjador per la IA subsimbòlica.

La IA explicable

En realitat, això de l'explicabilitat —paraula que m'atreveixo a proposar perquè encara no en tenim cap en català que em convenci prou per representar aquesta idea— no és cap nimietat. En els més de 60 anys de vida de la IA, hem passat d'uns primers algoritmes que es prenien tot el temps del món per resoldre la demostració d'un teorema amb una certa gràcia o per preocupar-se de si Kasparov podria guanyar o no la seva partida d'escacs, a una nova generació d'algoritmes que en pocs segons decideixen per tu quina pel·lícula t'agradarà més veure demà, si el banc et donarà o no el crèdit per comprar el pis, per quina carretera anem de casa a aquella muntanya meravellosa que hem vist a la televisió i ens han vingut ganes d'escalar, o si la senyora del 5è té una tomografia patològica i requereix urgentment un tractament oncològic... Sí, la

3 Un algoritme, o algorisme, és un conjunt de regles per a resoldre un problema —susceptible de ser expressat de forma matemàtica— en un nombre finit de passos. Per decisió de l'edició s'ha optat en tots els articles d'aquest dossier pel terme «algoritme».

4 Algoritme d'aprenentatge automàtic basat en una xarxa de nodes computacionals —neurones artificials— que processen ràpidament el senyal que els hi arriba i la propaguen a altres neurones, a imatge i semblança de com una xarxa de neurones cerebrals es passen informació per produir accions o pensaments. L'efecte conjunt de totes les neurones comunicant-se pot aprendre funcions matemàtiques de gran dificultat i elaborar models predictius en problemes complexos. Les prediccions podran ser bones, però és pràcticament impossible comprendre com s'han obtingut.

5 ROYAL SOCIETY, «Explainable AI. The basics».

El desenvolupament tecnològic vertiginós que hem viscut en la segona meitat del segle XX ha propiciat un progrés de la IA que ha permès avenços espectaculars

IA ha deixat enrere l'adolescència i ha començat a donar consells a tort i a dret a tothom qui li demana —i, el que és més preocupant, fins i tot a qui no li demana!—, en una escalada sense precedents des de la calculística complexa, a la total ingerència en l'àmbit de la presa de decisions.

Sens dubte el desenvolupament tecnològic vertiginós que hem viscut en la segona meitat del segle XX ha propiciat un progrés d'aquesta disciplina que ha permès avenços espectaculars pel que fa als temps de còmput, cada cop més curts, requerits per resoldre problemes sempre més i més complexos, més i més delicats, i de tots els àmbits esperables, des dels videojocs fins la política, passant per la salut, el medi ambient, la mobilitat o les finances. I és evident que en aquestes circumstàncies, si un algoritme ens ha de dir que mirem *Star Wars*, potser ens és ben igual com ho ha fet per endevinar que ens agradarà aquesta pel·lícula; total, si no l'encerta, amb no mirar-la, estem ben bé al cap del carrer... Però és clar, si li ha de dir a un alcalde —per exemple— que asfalti un carrer i no un altre, la cosa ja es complica, i comença a ser imprescindible que a més de recomanar què cal fer, ens pugui explicar el perquè de la decisió, oi? I vet-ho aquí que després d'una trajectòria de 64 anys, encara ens debatem entre la IA simbòlica, que s'explica tan bé però no escala —que vol dir que no pot resoldre problemes gaire grossos en temps curts— i la subsimbòlica, que no s'explica gens, però resol molt bé i molt de pressa problemes molt complicats. I així és com l'explicabilitat està a l'agenda de l'actual recerca en IA, i no va sola: l'acompanya una altra àrea emergent en aquest camp, l'«ètica» de la IA, que ens donaria per obrir un altre article igual de llarg que aquest, o potser més i tot.

El nucli dur: la cerca

Rebobinem una mica i analitzem millor el que ha passat en aquests 64 anys per ser on som ara.

La IA es gesta com la ciència de les activitats intel·ligents i creix al voltant de buscar solucions per problemes difícils. És una disciplina molt transversal, amb aplicacions a pràcticament tots els àmbits, però els problemes que són objecte de la IA comparteixen tots ells una característica intrínseca de la qual arrenca tota la seva complexitat, i alhora, tota la seva potència: la cerca. Els problemes que la IA resol involucren cerca, i una cerca no trivial. Cerca de la millor solució entre moltes, moltíssimes solucions possibles, que no es poden visitar totes fent força bruta per triar l'òptima —senzillament perquè n'hi ha massa. És més, gairebé m'atreviria a dir, que si un problema es pot resoldre sense implicar cerca, trobarem una altra ciència que identificarà una solució exacta, i probablement ho faci fins i tot millor que la que trobi la IA. És a dir, la IA estudia problemes on hi ha un nombre de solucions possibles enorme, molt combinatori, que no es poden calcular totes en temps finit, de tantes com són, i que a més, no sempre està molt clar quina és la millor. I així, els algoritmes propis de la IA se les enginyen per aplicar criteris que permetin estalviar-se una enorme part d'aquesta cerca i trobin solucions satisfactòries, raonables, potser subòptimes, al problema, en uns temps acceptables pel dia a dia de les decisions que anem prenent.

El *clustering* és només un dels moltíssims tipus de problemes dels que s'ocupa la IA

Posem un exemple en forma d'imatge molt plàstica per entendre-ho fàcilment. Classificar coses és una de les activitats més primàries de la ment humana. Tot el dia classifiquem de tot... culleres amb culleres i forquilles amb forquilles al calaix de la cuina, pantalons amb pantalons i jaquetes amb jaquetes a l'armari, bons clients amb bons clients, i els dolents amb els dolents, pacients sans amb sans, poc afectats amb poc afectats..., classifiquem de tot. De fet el *clustering* és la tècnica d'IA que resol el problema de trobar grups d'objectes semblants dins un conjunt d'objectes desendreçats —deixeu-m'ho dir així. I s'ocupa d'aquest problema des de fa molt temps, justament perquè respon a una d'aquestes activitats intel·ligents —en aquest cas intensiva— que la IA es proposa de modelar. Imaginem una persona d'aquestes que tenen una habitació feta una lleonera, que un bon dia decideix que vol endreçar la roba a l'armari. Què farà? El més probable és que primer faci grupets de peces, les camises a una banda, els pantalons a l'altra..., mitjons, jaquetes... I quan els tingui fets, busqui un espai per posar cada grup de coses a l'armari. Les camises penjades de la mitja barra, els pantalons de la barra llarga, les samarretes plegadetes en una lleixa... Què ha passat aquí? Teníem una pila de peces de roba barrejades i n'hem fet grups. I què faria una màquina que volgués endreçar l'armari? Doncs com que la pobra no sap res de roba, el que d'entrada es podria plantejar és anar muntant totes les formes possibles d'agrupar els objectes i anar mirant per cada agrupació si les classes que li surten són bones, és a dir, si surten classes que tenen només pantalons, o només camises. I anar provant... El que faria la màquina seria buscar entre totes les possibles combinacions, fins que trobés l'organització correcta. Quatre números a tall d'exemple: si tinguéssim l'habitació amb una barreja de cinc pantalons, cinc camises, cinc jaquetes, cinc samarretes i cinc *foulards*, hi hauria 10^{19} formes d'ordenar-les en cinc grups de coses! I totes aquestes ordenacions seran dolentes menys una! Totes tindran un pantaló —o més— al grup de les camises, i altres peces canviades de lloc, excepte una combinació. Només una de totes les combinacions possibles col·locarà tots els pantalons junts, totes les camises juntes, totes les jaquetes, etc. Només UNA d'entre totes les formes d'agrupar les 25 peces en cinc grups ens deixarà un armari ben endreçat quan assignem una zona de l'armari a cadascun dels grups fets. Això que un ésser humà fa en una estoneta un diumenge d'hivern a la tarda com aquell qui res —vencent la mandra, és clar—, a una màquina li costa 1.000.000.000.000.000.000.000 —deu trilions— de combinacions possibles! Dimensionem-ho? Suposem que la màquina triga un segon —moltíssim menys que un humà— per fer una agrupació en cinc grups de les 25 peces de roba... si la nostra IA hagués de provar totes les maneres d'agrupar les peces per trobar la solució bona —la que té tots els pantalons junts, totes les camises en un altre grup, i així...—, trigaria ni més ni menys que 3.170.979.198 de SEGLES! computant de forma interrompuda per poder endreçar el *ditxós* armari! I això que ordenar 25 peces de roba en cinc grups, no són res! En una botiga amb 100 peces de roba de tres classes diferents, hi ha... 10^{68} combinacions possibles per formar els tres grups! No hi ha vida que ho resolgui això... senzillament no és viable. Per això s'establiran criteris per escapçar aquesta cerca de manera que no haguem de formar totes les combinacions possibles per triar —escurçarem l'espai de cerca— i a més,

relaxarem les exigències, i si ens donen una agrupació amb alguna camisa posada entre els pantalons, mentre siguin poques les que estan canviades de grup, l'agrupació proposada també serà acceptable —no exigirem la millor solució, sinó el que s'anomena un *subòptim*, una de prou bona, però no caldrà que sigui la millor.

El *clustering* és només un dels moltíssims tipus de problemes dels que s'ocupa la IA. Però la majoria s'enfronta al problema que acabo d'exposar. El nombre de solucions possibles on buscar la solució bona és combinatori, i caldran criteris —anomenats *heurístics*— per navegar de forma estratègica per aquesta mar de solucions possibles i no visitar-les totes ni de bon tros, a canvi de pagar el preu que potser no trobarem ben bé la millor, sinó una de semblant, però ho farem en uns *tempos* acceptables.

La IA i les seves múltiples branques

La IA té moltíssimes branques, com el raonament automàtic —que sap encadenar peces de coneixement per viatjar d'unes premisses a unes conclusions seguint unes regles de raonament que poden ser més o menys sofisticades—; la planificació —el Google Maps per exemple, que sap enllaçar com anar per una xarxa viària, no en línia recta per damunt les cases, d'una geolocalització d'origen a una de destí—; l'aprenentatge automàtic —on situaríem el *clustering* del que parlàvem en l'anterior apartat, però també tots els algoritmes de reconeixement de classes, tan útils per tota labor de diagnòstic i majorment basats en buscar patrons de diferents formes sobre dades—; el processament del llenguatge natural —Google Translator, per exemple que és capaç d'entendre un text escrit i resumir-lo, traduir-lo o indexar-lo—; el reconeixement i/o generació de la parla —aquells bots que tant ens entretenen quan volem trucar a la companyia del llum per dir que tenim una avaria, oi?—; la visió per computador —capaç de trobar un tumor en una radiografia on l'ull humà no hi pesca res, entre moltes d'altres coses...—; la robòtica —on es treballa per exemple com identificar la força exacta per saber com un braç mecànic ha d'agafar un vas d'aigua per tal que no rellisqui i no esclati—; o els sistemes intel·ligents de suport a la presa de decisions —avui entre els més populars, incloent els recomanadors— que es nodreixen de moltes de les peces esmentades per donar suport a les decisions complexes. I de branques de la IA... Encara me'n deixo. Així doncs, l'àmbit de la IA és ampli, i profús. I dins aquest camp hi ha gent que treballa en coses molt diverses, des de fer cadires de rodes intel·ligents —que frenin soles si el conductor vol baixar una vorera massa alta i perilla—, fins a predir si plourà demà.

La IA, en popularitat ascendent!

De totes aquestes branques de la IA, n'hi ha una, l'aprenentatge automàtic, que posa la dada al centre de l'estructura i l'estudia per aprendre patrons de comportament sota diferents formes. Se situen en aquesta branca els mètodes de *clustering* —que endrecen armaris i tot el que troben... per entendre'ns—; els de diagnòstic —*classifiers* en anglès, però que no hem de confondre amb els que fan *clustering*—, els quals saben reconèixer per exemple si un pacient és sa o patològic i que diferencia l'un

El boom tecnològic que es deriva de les arquitectures paral·leles, l'arribada d'internet i l'aparició dels smart-sensors representa una revolució en les tècniques que treballen amb dades

de l'altre; i els predictius —que endevinen la facturació que podem esperar l'any que ve a l'empresa considerant les vendes que hem fet, la productivitat de la plantilla, i altres variables de context, per exemple, o la temperatura que farà demà en funció de les dades històriques... I dins d'aquests últims hi ha un tipus concret de mètodes, *les xarxes neuronals artificials*, que són mètodes connexionistes, i que simulant com transmeten la informació les xarxes de neurones del cervell, construeixen prediccions de les més complexes, superant per molt el paper que havia vingut fent la regressió —un mètode estadístic de finals del XIX, que es gesta en el context darwinià— que tan bon punt com s'enfronta a problemes on les relacions entre les variables no són directes, sinó una mica enrevessades, comença a comportar-se amb moltes deficiències. Per posar un exemple, l'espai que recorre un cotxe a velocitat constant és proporcional al temps que està en marxa. Aquest és un escenari on hi ha majorment —simplifiquem ara fregaments i altres forces que ens desvirtuen l'exemple— una relació lineal entre la velocitat i l'espai recorregut, i que es pot predir bé utilitzant models clàssics de regressió. Però, avui dia ens plantegem la predicció de fenòmens molt més complexes, com per exemple, el temps que trigarà una impressora 3D que funciona en una planta d'indústria 4.0 a parar-se per fallida d'algun component. De què depèn això? Hi intervenen molts elements, hi juguen les condicions de l'espai on està la impressora —temperatures ambient, humitats...— el tipus d'objectes que fabrica, els paràmetres de fabricació, l'estat del material que utilitza per imprimir, com la manipula l'operari... I a més, la influència de tots aquests factors en aquest temps que volem predir... no és directa!

Les xarxes neuronals artificials, proposades per primer cop per Warren McCulloch (1898-1969) als anys 1940 sí que poden modelar relacions difícils,[6] i ho fan molt bé, però s'han d'entrenar primer, amb dades històriques que permeten aprendre els patrons del problema que s'estudia... i resulta que per estabilitzar-les, necessiten moltes més dades de les disponibles en aquella època, i una potència de càlcul molt superior a la del moment. Si bé eren una eina prometedora —de caixa negra, sí, no explicable, però capaç de fer prediccions de fenòmens força més complicats que el seu paral·lel estadístic— requerien d'una infraestructura tecnològica prohibitiva pel moment històric. Així que a la dècada de 1970 queden estancades i cauen més o menys en l'oblit. El boom tecnològic que es deriva de les arquitectures paral·leles, l'arribada d'internet i l'aparició dels *smart-sensors* representa una revolució en les tècniques que treballen amb dades, primer perquè abarateixen moltíssim la recol·lecció i l'enviament de dades a altres llocs, i en segon lloc perquè es redueixen de forma molt rellevant els *tempos* de computació i permeten plantejar-se prediccions de fenòmens de tal complexitat, que mai abans s'havien pogut estudiar. És a partir dels anys 1980 on ressorgeixen aquestes tècniques i es comencen a utilitzar, ja no sobre les dades habituals sinó per a processar imatges —i més tard textos— amb un impacte important sobre els mètodes de diagnosi basats en imatge mèdica.

6 MCCULLOCH; PITTS, «A Logical Calculus of Ideas Immanent in Nervous Activity».

L'ús i abús de la dada, no sempre tan contingut ni respectuós com seria d'esperar, ha generat una problemàtica de molt difícil solució i que actualment està centrant el debat de la IA

Beneïda dada, font de tota santedat

A finals de la dècada de 1990 s'agafa plena consciència del valor que té la dada en els processos de presa de decisions. La cadena de supermercats Walmart, esbrinant quin era el perfil del petit percentatge de clients que li compraven bolquers i cerveses els divendres a la tarda, va poder dissenyar una estratègia dirigida que li va incrementar les vendes molt significativament. Aquest llegendari cas va posar al centre de l'huracà l'explotació de les dades per establir les estratègies de negoci. Bé, honestament, del valor de la informació per les decisions ja se'n parlava als anys 1960, com recull Hans Peters Luhn (1896-1964), entre d'altres.[7] El que passa és que en aquella època, la informació depenia més del coneixement del sector que tenia el directiu, que no pas de dades, encara escasses i costoses d'obtenir. No serà fins que hi ha les condicions idònies, que neix el que s'anomena «mineria de dades», que més tard evolucionarà al que avui es coneix com «ciència de les dades»[8] on, simplificant molt, la IA, l'estadística i la computació col·laboren per portar la dada als processos de decisió, i ja no només a àmbits com la medicina o els relacionats amb el medi ambient, sinó de ple al cor del món dels negocis. Sempre m'ha impactat que Hal Varian (1947), economista en cap de Google al 2009 es queixés que «*data are widely available; what is scarce is the ability to extract wisdom from them*»[9] és a dir, que de dades en teníem muntanyes i el que no hi havia era persones formades per extreure'n el valor que contenien.[10] Mentrestant, Craig Mundie (1949), assessor directiu de Microsoft, vaticinava al 2010 que venia el que ell anomenava la «*data-centered economy*», l'economia centrada en la dada. I tant que venia! Al 2020 s'assumeix que el valor de les organitzacions ja no és la tecnologia, ni el capital, ni la maquinària, ni les patents de fabricació, ni tan sols el capital humà, ni el *know-how*... sinó «la dada». I així resulta que actualment tothom recull dades de tot i de per tot arreu i mira de modelar-les utilitzant tècniques d'IA per poder fer models predictius que, pel que sembla, estan al centre de tots els negocis haguts i per haver, i no només dels negocis. Actualment s'estan recollint dades de quasi tot el que és susceptible de generar-ne, especialment el mòbil, els rellotges, els nostres comptes web, de Google... Estem envoltats de «tercers ulls virtuals» per tot arreu, que tot ho veuen i tot ho xerren, amb més o menys permís per part nostra. I així ens tornem a trobar davant del dilema ètic que planteja la IA —i que com ja he dit abans, deixaré a part en aquest article. En definitiva, és cert: l'ús i abús de la dada, no sempre tan respectuós ni contingut com seria d'esperar, ha generat una problemàtica de molt difícil solució i que actualment està centrant el debat de la IA a casa nostra, però també a Europa i altres indrets.

Parlant d'abusos, podríem també esmentar l'abús de llenguatge que ens envolta. En efecte, avui dia resulta confús què es vol dir quan es declara que s'utilitza IA en una empresa. Què hi ha realment al darrere? Ciència de les dades o aprenentatge automàtic, o només una etiqueta de moda que

7 LUHN, *A Business Intelligence System*.

8 GIBERT, «Special Issue on Environmental Data Science. Applications to Air quality and Water cycle».

9 En català: «Les dades estan àmpliament disponibles; el que escasseja és la capacitat d'extreure'n saviesa» [traducció de l'edició].

10 PALOS, «The information flow in virtual teams».

Si volem donar suport a les decisions en la IA necessitem models explicables, que justifiquin llurs resultats de forma comprensible

s'utilitza perquè fa bonic i millora substancialment el posicionament de mercat? I aquesta cara més qüestionable de la IA, conviu amb la seva oposada: la de totes les oportunitats i noves capacitats que la IA pot generar per curar malalties greus, lluitar contra el canvi climàtic, i tantes altres aplicacions ètiques que poden fer un món millor.

Última parada... el *deep learning*

A aquesta nova concepció on el que té valor és la dada, s'hi ha d'afegir un últim avenç significatiu en l'àmbit de les xarxes neuronals: el *deep learning* [aprenentatge profund].[11] Un tipus especial de xarxes neuronals que han aconseguit reduir els temps de computació de forma escandalosa, fent viable l'ús de dades massives a les organitzacions, sense més limitació que la pròpia ètica i escrúpol d'aquell qui recull la dada i l'explota. Però en realitat és aquesta la clau de volta. Perquè ha estat el *deep learning*, el germà petit de les xarxes neuronals i de l'aprenentatge automàtic, però el més espavilat de tots, el que ha escalat la IA als nivells productius competitius que han permès incorporar-la al teixit empresarial de forma massiva, transformant així el cor de les organitzacions de forma important.

I així, cap a on anem?

Disposem actualment de les més avançades tecnologies per aprendre dels fenòmens més difícils, no només gràcies al *deep learning*, sinó també amb altres mètodes d'aprenentatge automàtic, potser no tan escalables, però molts cops més explicables. I la seva aplicabilitat és clara: disposar de models que ens assisteixen a l'hora de prendre decisions difícils. En l'actual escenari, doncs, envoltats com estem ja d'IAs que suggereixen, fan i desfan cada dia, i amb perspectives que en vinguin més a acompanyar la nostra convivència, esdevé crític desenvolupar la cultura necessària per a fer-ne un bon ús. La realitat és que, sovint, tenir moltes més dades no és garantia d'aprendre millor del món ni de tenir millors prediccions. Hi ha un greu problema per detectar la qualitat de les dades, sovint esbiaixades per construcció, i encara no tenim criteris sòlids per determinar ni mesurar aquest biaix, que indefectiblement es perpetua en les prediccions que se'n derivin, devaluant el valor d'aquests models, i de conseqüències incertes quan són la base per definir estratègies o prendre decisions. Si volem donar suport a decisions en la IA, necessitem models explicables, que justifiquin llurs resultats de forma comprensible. D'altra banda, la petjada ambiental de la IA és preocupant; o millor dit, no tant de la IA com de la dada: caldrien 2,5 milions d'aerogeneradors al món només per mantenir funcionant els 500.000 *data-centers* que hi ha al planeta.[12] La IA és una tecnologia que pot portar grans millores per la vida de les persones en la quotidianitat a nivell mundial, però ve acompanyada d'algunes incerteses que convé no obviar. Els usuaris han de poder entendre què fan les aplicacions, quines dades consumeixen, quines conseqüències tenen i, si és el cas, quins drets vulneren o podrien vulnerar, per tal de poder triar com es relacionen amb les IA que els envolten.

11 HINTON; OSINDERO; TEH, «A fast learning algorithm for deep belief nets».

12 POGANČIĆ, «The Carbon Footprint of IA Research».

Fem-ho, però fem-ho bé!

La IA ha escalat al teixit productiu i té un enorme potencial per al desenvolupament del país. Catalunya té una concentració de talent en IA extraordinària, amb una de les comunitats científiques més cohesionades d'Europa[13] i un ecosistema de gran valor; i fou, també, pionera en advertir del risc, en la «Declaració de Barcelona de la IA»[14] de 2017, de deriva negativa en els usos de la IA. La IA representa una salt tecnològic, i com tot salt conté incerteses i requereix algunes precaucions. La IA ha deixat de ser un mite i avui és una realitat quotidiana. La IA obre la porta a envoltar-nos de moltes aplicacions que contribueixin a construir un món més just. Si en fem bon ús, en traurem molts beneficis. Aprofitem-ho! El repte s'ho val. ■

13 GIBERT, «La Intel·ligència Artificial a Catalunya, una visió personal».

14 STEELS; LÓPEZ DE MÁNTARAS, «The Barcelona declaration for the proper development and usage of artificial intelligence in Europe».

Bibliografia

GIBERT, Karina. «La intel·ligència Artificial a Catalunya, una visió personal». A *IDEES, revista de les idees*, núm. 48. Centre d'Estudis de Temes Contemporanis. Barcelona: Generalitat de Catalunya, 2020.

GIBERT, Karina (et al.). «Special Issue on Environmental Data Science. Applications to Air quality and Water cycle». A *Environmental data science. Environmental Modelling & Software*, volum 106, pàg. 4-12, 2018.

HINTON, Geoffrey; OSINDERO, Simon; TEH, Yee-Whye. «A fast learning algorithm for deep belief nets» [en línia]. A *Neural computation*, núm. 18(7), pàg. 1527-1554. Disponible a: <www.cs.toronto.edu>.

LUHN, Hans Peter. ***A Business Intelligence System*** [en línia]. A *IBM Technical Journals*, 1958. Disponible a: <www.altaplana.com>.

McCARTHY, John (et al.). «A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence» [en línia]. Disponible a: <www.jmc.stanford.edu>.

MCCULLOCH, Warren; PITTS, Walter. «A Logical Calculus of Ideas Immanent in Nervous Activity». A *Bulletin of Mathematical Biophysics*, núm. 5 (4), pàg. 115-133, 1943.

MINSKY, Marvin. ***Heuristic aspects of the artificial intelligence problem.*** Virginia: Services Technical Information Agency, 1956.

PALOS, Andra. «The information flow in virtual teams» [en línia]. A *Journal of Advanced Research in Management*, núm. 3(05), pàg. 38-45, 2012. Disponible a: <www.deepdyve.com>.

PALOS, Andra. «The information flow in virtual teams» [en línia]. A *Journal of Advanced Research in Management*, núm. 3(05), pàg. 38-45, 2012. Disponible a: <www.deepdyve.com>.

POGANČIĆ, Marin Vlastelica. «The Carbon Footprint of IA Research» [en línia]. A *Towards data science*, a 1 d'octubre de 2019. Disponible a: <www.towardsdatascience.com>.

ROYAL SOCIETY. «Explainable AI. The basics» [en línia]. Disponible a: <www.royalsociety.org>.

STEELS, Luc; LOPEZ DE MÁNTARAS, Ramon. «Barcelona Declaration for the proper development and usage of artificial intelligence in Europe» [en línia]. Disponible a: <www.content.iospress.com>.