

Sinapsi de l'èsser

Una neurona artificial dubtosa

Francisco Grimaldo, Emilio Soria, Emilia López (ETSE UV)

Vos proposem fer una passejada pels treballs del segle passat sobre com modelitzar el funcionament del cervell humà, que són la base de les aplicacions més precises de reconeixement d'imatges, els assistents personals com ara Siri o Cortana, els traductors automàtics o els cotxes autònoms.

Les dades han arribat per a quedar-se i estan canviant la manera en què percebem i prenem decisions en àrees com la sanitat, les relacions socials, la política, el medi ambient i els negocis. La ciència de dades està pertot i ens permet adquirir, processar i extraure coneixement a partir de grans volums de dades mitjançant tècniques d'avantguarda heretades del món de les matemàtiques, l'estadística, la informàtica, la intel·ligència artificial, l'economia i la salut.

La gran potència dels ordinadors i l'enorme capacitat dels nous dispositius de memòria han fet possible la utilització de models d'aprenentatge automàtic que tenen una alta càrrega computacional i que, fins fa uns anys, eren impossibles d'aplicar. Entre aquests models destaquen les xarxes neuronals artificials o models connexionistes. Dins d'aquesta categoria, l'aprenentatge profund (deep learning) està de moda i constitueix l'estat de l'art de la majoria dels sistemes experts actuals. Les aplicacions més precises de reconeixement d'imatges, els assistents personals com ara Siri o Cortana, els traductors automàtics, els cotxes autònoms... es basen sovint en models neuronals.

Les dendrites d'aquestes neurones, però, s'arrelen en la saviesa que desprenen els treballs del segle passat sobre com modelitzar el funcionament del cervell humà. Així, doncs, vos

proposem fer una passejada per aquest cementeri de neurones oblidades, tot destacant els fonaments matemàtics i descrivint els contextos d'aplicació en què han triomfat.

Podem fixar el punt de partida en l'estudi del cervell a finals del segle XIX, quan Santiago Ramon y Cajal ocupava la càtedra d'anatomia de la Universitat de València. És a València que Cajal va encetar el conreu de la seua investigació capdavantera en neurohistologia i, des de Barcelona estant, que va presentar a la comunitat científica, en el 1r Congrés Mèdic-Farmacèutic Regional de València, la seua llei de la polarització dinàmica que explica el funcionament del component més xicotet en l'estructura del cervell: la neurona.

Al llarg de la primera meitat del segle XX, el psicòleg Karl Lashley va estudiar la memòria i l'aprenentatge per concloure que eren processos distribuïts, no localitzats en una àrea determinada del cervell. Un estudiant seu, Donald Hebb, en va recollir el testimoni i va determinar una de les regles referides més comunament pel connexionisme: l'aprenentatge hebbià. El seu llibre *The Organization of the Behavior* especifica per primera vegada una regla per a la modificació de les sinapsis, açò és, una regla fisiològica d'aprenentatge. A més a més, proposa que la connectivitat del cervell canvia contínuament a mesura

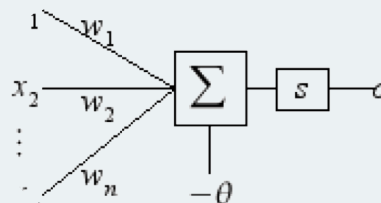
que un organisme aprèn coses noves, tot creant associacions neuronals amb aquests canvis.

En el seu postulat de l'aprenentatge, Hebb segueix allò suggerit per Cajal quan afirmava que l'efectivitat d'una sinapsi variable entre dues neurones s'incrementa per una activació repetida d'una neurona sobre l'altra a través d'aquesta sinapsi. Des d'un punt de vista neurofisiològic, la regla plantejada per Hebb és una regla variant-temporal, amb un alt mecanisme interactiu que incrementa l'eficàcia sinàptica com una funció de l'activitat pre i post sinàptica. Des d'un punt de vista connexionista, la regla de Hebb és un tipus d'aprenentatge no supervisat (que no necessita de cap «mestre») en el qual les connexions entre dues neurones s'incrementen si ambdues s'activen al mateix temps.

les característiques de funcionament del que, posteriorment, es coneixerà com la neurona de McCulloch-Pitts. Aquest model de neurona representa una mena de dispositiu binari amb dos possibles estats d'eixida: actiu o inactiu (1 o 0). Llavors, la neurona té un llindar de funcionament per sota del qual està inactiva i pot rebre entrades inhibidores o excitadores que, respectivament, la mantinguen dreta o fan que superi aquest llindar, el que produirà la seua activació.

El model de McCulloch-Pitts s'apropava a allò que la neurofisiologia coneixia el 1943 al voltant de l'activitat sinàptica neuronal i, a més a més, permetia sintetitzar alguna de les funcions lògiques de l'àlgebra booleana que suportaven els fonaments de la computació. Aquesta capacitat va desencadenar l'eufòria

Anatomia matemàtica d'una neurona artificial



Per a calcular l'estat d'eixida (o), la neurona agrega el llindar (θ) a les entrades (x_k) en una combinació lineal que fa servir els pesos de les sinapsis (w_k). Darrerament, han anat sorgint variacions que realitzen una transformació quadràtica o polinòmica de l'entrada.

La funció d'activació en la neurona de McCulloch-Pitts és la funció llindar:

$$s(x) = \begin{cases} 1 & x \geq 0 \\ 0 & x < 0 \end{cases}$$

En l'època daurada del 80, es va fer servir la sigmoide com a funció d'activació i, hui dia, la funció ReLU ($\max(\Sigma, 0)$) es la més usada, ja que és molt fàcil de calcular i no té problemes amb el descens del gradient. El model dibuixat és estàtic però, quan l'eixida d'una neurona es retroalimenta a una de les seues entrades, la neurona es comportarà de manera dinàmica i tindrà memòria, ja que el seu estat d'activació en un cert instant dependrà del seu estat anterior i de les altres entrades externes.

La següent gran contribució a considerar és el treball de Warren McCulloch i Walter Pitts, que hom sol identificar com l'inici de la curta però intensa història de les xarxes neuronals artificials. Aquest treball va definir

en l'ús de les neurones artificials així com en els somnis científics. Si podíem reproduir el comportament d'una neurona individual, per què no podríem implementar tot un sistema de coneixement mitjançant l'ús d'un

conjunt d'aquestes neurones artificials?

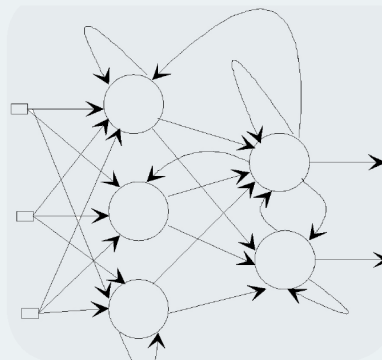
Guiat per aquest miratge, Alan Turing va ser un dels precursors del connectivisme allà pel 1948, encara que el seu article *Intelligent Machinery* no va veure la llum perquè havia estat titllat d'assaig escolar pel seu cap, Sir Charles Darwin (nét del famós naturalista de qui suposem no havia heretat una certa miopia científica). Un altre geni de les matemàtiques i pare de la informàtica, John Von Neumann, també va suggerir que l'estudi del sistema nerviós central i de les idees connexionistes era un possible camí per a millorar els ordinadors.

pel que l'autor anomenava «dimonis». La segona, el Perceptró, proposava una nova aproximació al problema del reconeixement de patrons i va atraure irresistiblement els enginyers i les enginyeres, pel fet que representava una màquina capaç d'aprendre. Com a conseqüència, el 1960, Bernie Widrow i Ted Hoff van presentar el seu ADALINE, un sistema adaptatiu que podia aprendre de forma més precisa i ràpida que els perceptrons existents.

La desfeta va arribar a tocar dels anys 70, quan Marvin Minsky i Seymour Papert van publicar el seu article *Perceptrons*. Aquest document

Arquitectures neuronals artificials

És habitual que els elements d'una xarxa neuronal artificial s'interconnecten seguint una disposició definida per capes. Quan existeixen enllaços de retroalimentació entre capes o dins de la mateixa capa, ens trobem davant d'estructures recurrents.



Xarxa multicapa recurrent

En l'època daurada del 80, es va fer servir la sigmoide com a funció d'activació i, hui dia, la funció ReLU ($\max(\Sigma, 0)$) es la més usada, ja que és molt fàcil de calcular i no té problemes amb el descens del gradient. El model dibuixat és estàtic però, quan l'eixida d'una neurona es retroalimenta a una de les seues entrades, la neurona es comportarà de manera dinàmica i tindrà memòria, ja que el seu estat d'activació en un cert instant dependrà del seu estat anterior i de les altres entrades externes.

Així, els investigadors N. Rochester, J. Holland, L. Haibt i W. Duda van verificar el 1956 una teoria neuronal basada en el postulat de Hebb i, en les darreries de la dècada dels 50, Oliver Selfridge i Frank Rosenblatt van reeixir amb el disseny dues arquitectures connexionistes. La primera, el Pandemonium, constava d'una sèrie de capes que es repartien les diferents tasques a realitzar i estaven formades

posava de manifest que els perceptrons de l'època només podien resoldre problemes linealment separables dels quals, dissortadament, hi havia ben pocs. El treball de Minsky i Papert va paraitzar durant 10 anys l'avançament d'aquest camp de la intel·ligència artificial, però alguns grupuscles hi continuaren treballant. La resistència preparava el seu retorn...

En els primers anys obscurs, Teuvo Kohonen i James Anderson van proposar el mateix model de memòria associativa de forma simultània. Aquesta coincidència fa palès fins a quin punt els sistemes connexionistes son multidisciplinaris, ja que ambdós autors tenen una formació molt diversa: Kohonen és enginyer elèctric i Anderson és psicòleg i neurocientífic. En el model artificial plantejat, la neurona és un sistema lineal que empra com a regla d'aprenentatge la regla de Hebb modificada d'acord amb un mecanisme d'associació lineal: el canvi en la sinapsi és proporcional al producte entre l'entrada i l'eixida de la neurona.

resoldre problemes d'optimització (p. ex. el clàssic problema del viatjant). El mateix any 1982, Kohonen va publicar un important article sobre mapes auto-organitzatius que s'ordenaven d'acord amb unes regles senzilles i que aprenien automàticament sense conèixer les solucions als casos d'exemple utilitzats durant la fase d'entrenament de la xarxa, açò és, de manera no supervisada.

L'any següent, en el número especial sobre models neuronals de la revista IEEE Transactions on Systems, Man and Cybernetics, apareixen dos treballs de relleu per al desenvolupament de les xarxes neuronals. Kunihiro

Arquitectures neuronals artificials

Les xarxes neuronals artificials fan servir uns procediments d'entrenament el qual, partint d'uns casos d'exemple, permeten ajustar automàticament els pesos de les sinapsis per a proporcionar l'eixida desitjada: l'algorisme d'aprenentatge.

En els algorismes d'aprenentatge no supervisats el protocol no necessita conèixer el resultat que ha de donar la xarxa per a cada cas d'exemple. La xarxa és capaç d'auto-organitzar-se a través de l'agrupament, segons les seues característiques, dels diferents senyals d'entrada.

L'aprenentatge supervisat, pel contrari, proporciona a la xarxa la resposta correcta als exemples que se li presenten en la fase d'entrenament. Llavors, la xarxa modifica els pesos d'acord amb la diferència entre la seua eixida i el senyal desitjat.

En l'aprenentatge per reforç, la xarxa ajusta els seus pesos sabent només si la seua eixida coincideix amb la buscada; açò és, la informació per a l'aprenentatge és de tipus booleà (vertader o fals). En l'aprenentatge per correcció, la magnitud de l'error comès determina la magnitud en el canvi dels sinapsis.

El 1980, Stephen Grossberg, un dels autors més prolífics en el camp de les xarxes neuronals artificials, va crear un nou principi d'auto-organització gràcies a les xarxes neuronals ART (Adaptive Resonance Theory). Però no és fins 1982, que el distingit físic teòric John Hopfield publica un treball clau per al ressorgiment del camp. Hopfield va desenvolupar la idea de l'ús d'una funció d'energia per a comprendre la dinàmica d'una xarxa neuronal recurrent, en què l'estat d'una neurona depèn d'altres estats precedents. El principal ús d'aquestes xarxes ha estat, doncs, com a memòries i com a instrument per a

Fukushima, Sei Miyake i Takayuki Ito presenten el Neocognitron, que combina idees del camp de la fisiologia, de l'enginyeria i de la teoria neuronal per a crear un dispositiu que és capaç de reconèixer nombres escrits a mà. El segon treball, presentat per Andrew Barto, Richard Sutton i Charles Anderson, mostra com es pot controlar el balanceig d'una corriola mitjançant l'aprenentatge per reforç, en què només cal conèixer el signe del error que s'ha comès per a poder aprendre.

El 1986 apareix un treball que, junt amb el de Hopfield, ressuscitarà l'interès per les xarxes neuronals. En aquest treball,

David Rumelhart, Geoffrey Hinton i Ronald Williams desenvolupen l'algorisme d'aprenentatge per retropropagació (backpropagation) per a xarxes neuronals multicapa i donen una sèrie d'exemples en què mostren la potència del mètode desenvolupat. Val a dir que, tot i que aquest treball va suposar la publicitat mundial del procediment de retropropagació, Paul Werbos ja l'havia descrit 12 anys abans en plena horabaixa.

Des d'aleshores, el nombre de treballs sobre xarxes neuronals va augmentar exponencialment fins a arribar a la meitat de la dècada dels 90, quan va aparèixer un model d'aprenentatge màquina que rivalitzava en prestacions amb els models neuronals: les màquines de vector suport. Tanmateix, dos articles influents de Hinton en 2006 van fer renàixer definitivament el camp de les xarxes neuronals de forma espectacular i el van capbussar en el que es coneix com aprenentatge profund o deep learning.

Les xarxes neuronals profundes es troben al nucli de més del 90% dels sistemes experts implementats en l'actualitat, fonamentalment, perquè milloren els resultats d'altres models matemàtics quan el problema a tractar presenta les següents característiques:

1. Són problemes difícils de resoldre on les relacions que hom tracta de trobar són complexes (no linealitat).
2. Les dades són imprecises o contenen

pertorbacions estadístiques (presència de soroll).

3. El problema necessita un gran nombre de variables dependents per a ser definit (alta dimensionalitat).

4. L'entorn és variable, açò és, les relacions entre les variables canvien amb el temps (problemes dinàmics).

Així, doncs, aquests models s'han emprat sobre manera en totes les àrees del coneixement, entre les quals podem destacar les següents aplicacions:

Medicina: pronòstic de cardiopaties a partir de l'electrocardiograma; detecció de tumors cancerígens fent servir imatges; compressió de senyals bioelèctrics per a la diagnòsi remota; predicció de malalties degeneratives, etc..

Farmàcia: predicció del risc d'intoxicació per l'administració d'un fàrmac; anàlisi de la resposta emètica en tractaments amb quimioteràpia; anàlisi dels efectes adversos de la concentració d'un fàrmac en la sang; etc.

Tecnologies de la informació i les comunicacions: reconeixement de la veu en sistemes d'assistència; detecció d'objectes en imatges de sonar i radar; recuperació de la informació rebuda per un canal de comunicació amb distorsió; identificació de patrons d'atac en contextos de ciberseguretat, etc.



Economia: determinació del factor de risc en concessions creditícies; detecció de fraus en l'ús de targetes de crèdit; estimació del risc de fallida d'un banc; predicció de la despesa elèctrica d'empreses; previsió de la necessitat d'inventari d'un negoci, etc.

Medi ambient: anàlisi de la pol·lució atmosfèrica en zones urbanes i rurals; predicció d'irradiació solar; previsió de catàstrofes naturals en funció de les variacions globals de la temperatura; cancel·lació activa de la contaminació acústica; control automàtic l'abocament de residus, etc.

Siga quina siga l'aplicació i el seu grau de complexitat, la bellesa de les xarxes neuronals artificials es troba en la senzillesa dels seus components estructurals. La genialitat del comportament dels models connexionistes rau, no en la capacitat individual de les neurones artificials, sinó en la riquesa de les seues interaccions i en la seua capacitat de combinació amb altres tècniques.

La investigació en computació quàntica, que encara està donant les seues primeres passes, n'és un bon exemple. Les xarxes neuronals quàntiques malden per mantindre la màgia tot i substituir l'eixida binària de la neurona de McCulloch-Pitts per un qubit (unitat de mesura quàntica de la informació). Aquesta aproximació permetria obrir nous horitzons liderats per un component l'eixida del qual podria ser, alhora, una superposició dels seus estats actiu i inactiu: una neurona artificial dubtosa!

Això, però, es una altra història... Mentrestant, continuarem escrivint i llegint versos binaris com: «*Aquesta cançó és un bes, per a despertar, a Alan Turing*»¹.

¹ Tornada de la cançó "El bes" de l'àlbum "Un dígit binari dubtós", recital per a Alan Turing tret per Hidrogenesse l'any 2012.

