YANN LE CUN

A mesterséges intelligencia és a mélytanulás forradalma

YANN LE CUN

A mesterséges intelligencia és a mélytanulás forradalma

FORDÍTOTTA
MOLDVAY TAMÁS

TYPOTEX

Cet ouvrage a bénéficié du soutien des programmes d'aide
la traduction d'ouvrages français en langues étrangères de CNL.

A mű eredeti címe:
Quand la machine apprend. La révolution des neurones artificiels et de l'apprentissage profond.

Lektorálta Kertész Gábor

ISBN 978 963 493 335 9

TARTALOM

BEVEZETÉS

„HAL, nyisd ki a légzsilip ajtaját!” A *2001: Űrodüsszeia* című filmben az űrhajó küldetését irányító szuperintelligens számítógép, a HAL9000 nem hajlandó kinyitni a légzsilipet Dave Bowman űrhajós előtt. A jelenet drámaian összefoglalja a mesterséges intelligencia problematikáját. A rendszer az őt megalkotó ember ellen fordul. Vajon ez csak fantázia, vagy jogos félelem? Kell-e attól tartanunk, hogy világunkat egy napon terminátorok, kvázi korlátlan hatalommal és sötét szándékkal fellépő humanoid robotok vetik uralmuk alá? Napjainkban egyre gyakrabban teszik fel el ezt a kérdést, hiszen jelenleg egy olyan forradalmat élünk át, amelyet ötven évvel ezelőtt el sem lehetett volna képzelni. A mesterséges intelligencia, amelynek kutatói karrieremet szenteltem, gyökerestül felforgatja társadalmunkat.

Ezt a könyvet azért akartam megírni, hogy elmagyarázzam az ezzel kapcsolatos módszerek és technikák összességét, anélkül hogy elleplezném a bonyolultságukat. Ez a vállalkozás kevésbé egyszerű, mint megtanítani valakit dámajátékot játszani, de úgy gondolom, mégiscsak szükséges belevágni, hogy az emberek megfontolt véleményt alkothassanak a dologról. A média tele van a „mélytanulás” *(deep learning),* „gépi tanulás” *(machine learning),* „neurális hálózatok” kifejezésekkel… Lépésről lépésre szeretném megvilágítani a mögöttük meghúzódó tudományos eljárást, amely a számítástechnika és az idegtudomány kereszteződésében helyezkedik el. Próbálom ráadásul mindezt úgy tenni, hogy ne puszta metaforákban beszéljek.

Utazásunk a gépezet kellős közepébe az olvasás két szintjét kínálja fel az olvasónak. Az első szint intuitívabb. Mesélek, leírok, elemzek. Aztán időről időre haladóbb matematikai és informatikai fejtegeté-

sekbe bocsátkozom, amelyeket a téma iránt mélyebben érdeklődő olvasóknak szánok.

A mesterséges intelligencia (MI) lehetővé teszi, hogy egy gép felismerjen egy képet, átültessen egy beszédet egyik nyelvről a másikra, lefordítson egy szöveget, automatizálja az autóvezetést vagy egy ipari folyamat irányítását. Az elmúlt években tapasztalt bámulatos haladás a *deep learning*nek köszönhető, amely lehetővé teszi, hogy a gépet konkrét beprogramozás helyett megtanítsuk arra, hogyan hajtsa végre a feladatot. A *deep learning* a mesterséges neurális hálózat sajátossága, melynek architektúráját és működését az agy ihlette.

Az emberi agy átlagosan 86 milliárd neuronból áll, és ezek az idegsejtek egymással össze vannak kapcsolva. A mesterséges neurális hálózatok is sok egységből, sok matematikai függvényből állnak, amelyek nagyon leegyszerűsítve a neuronokhoz hasonlítanak. Az agyban a tanulás megváltoztatja az idegsejtek közötti kapcsolatokat, és ugyanez igaz a mesterséges neurális hálózatokra is. Mivel ezek az egységek gyakran több réteget alkotnak, ezért „mély" hálózatokról és „mély" tanulásról beszélünk.

A mesterséges neuronoknak az a feladatuk, hogy kiszámítsák a bemeneti jeleik súlyozott összegét, és amennyiben ez az összeg egy bizonyos küszöbértéket meghalad, kimenő jelet állítsanak elő. A mesterséges neuron tehát se több, se kevesebb, mint egy számítógépes program által kiszámított matematikai függvény. Persze nem véletlen, hogy a mesterséges intelligencia területének szóhasználata közel van az agyéhoz, hiszen az idegtudomány felfedezései ösztönözték a mesterséges intelligencia kutatását.

A jelen könyvben szeretném végigkövetni a saját szellemi utazásomat is ebben a rendkívüli tudományos kalandban. A nevem összeforrott az ún. „konvolúciós" hálózatokkal, amelyek átalakították a vizuális felismerést. Az emlősök látókérgének szerkezete és működése ihlette konvolúciós hálózatok lehetővé teszik a kép, videó, hang, beszéd, szöveg és egyéb jeltípusok hatékony feldolgozását.

Mit csinál egy kutató? Honnan jönnek az ötletei? Én a magam részéről sokat dolgozom az intuícióra támaszkodva. A matek ezután következik. Tudom, más tudósok pont fordítva működnek. Én határeseteket vetítek ki a fejemben, amiket Einstein *„gedanken experimentnek"*, azaz *„gondolatkísérletnek"* nevezett: elképzelünk egy helyzetet, majd megpróbáljuk mérlegelni a következményeit, hogy jobban megértsük a problémát.

Ezt az intuíciót az olvasmányaim táplálják. Valósággal faltam a könyveket, elmerültem az előttem járók munkájában. Soha nem egyedül teszünk felfedezéseket. Az ötletek ott vannak, ott szunnyadnak, mígnem egyikünk-másikunk fejében felébrednek, mert elérkezett az idő. Így folyik a kutatómunka. Kicsit összevissza halad előre, ugrál, helyben topog… olykor még hátra is lép. De mindig kollektív. A laboratóriumában magányosan dolgozó feltaláló képe romantikus fikció.

A mélytanulás kalandja nem volt mentes a nehézségektől. Minden oldalról meg kellett küzdenünk a szkeptikusokkal. A kizárólag a logikán és a kézzel írt programokon alapuló mesterséges intelligencia hívei bizton állították, hogy kudarcot fogunk vallani. A „klasszikus" gépi tanulás bajnokai ujjal mutogattak ránk. Az általunk kutatott mélytanulás mindazonáltal csupán néhány sajátos technikára terjedt ki a gépi tanulás tágabb területén belül. Csakhogy az utóbbinak, amely lehetővé teszi, hogy a gép példákon keresztül tanuljon meg egy feladatot, anélkül hogy kifejezetten erre programozták volna, megvoltak a maga korlátai. Mi próbáltuk kitolni ezeket. A mély neurális hálózat, az általunk javasolt *deep learning* volt ehhez az eszköz. Nagyon hatékonyak voltak, de komplikált volt működésre bírni őket, és nehéz volt a matematikai elemzésük. Szóval afféle alkimistának számítottunk…

A „klasszikus" gépi tanulás hívei 2010 körül hagytak fel a neurális hálózatokon való gúnyolódással, amikor azok végre bebizonyították hatékonyságukat. Én a magam részéről soha nem kételkedtem bennük. Mindig is meg voltam győződve arról, hogy az emberi intelligencia annyira komplex, hogy ahhoz, hogy leutánozzuk, egy olyan önszervező rendszer megalkotását kell megcéloznunk, amely képes önmagától, a saját tapasztalataiból tanulni. Ma már egyértelműen a mesterséges intelligenciának ez a formája a legígéretesebb, és a fejlesztését a rendelkezésünkre álló nagy adatbázisok és az olyan eszközök is elősegítik, mint a GPU (*graphical processing unit,* azaz grafikai processzor vagy más néven grafikus gyorsító), amelyek megtöbbszörözik a számítógépek számítási teljesítményét.

Franciaországi tanulmányaim végén azt terveztem, hogy egy-két évet Észak-Amerikában töltök. De végleg kint ragadtam! Utam során, sok-sok kaland után csatlakoztam a Facebookhoz, a 2 milliárd előfizetővel rendelkező közösségi platformhoz, hogy ott irányítsam az MI-vel kapcsolatos alapkutatást. A Facebooknál is nyitott dosszi-

ékkal dolgozom. Nem akarok eltitkolni semmit azzal kapcsolatban, hogy mi folyik Mark Zuckerberg 2018-ban súlyosan meggyanúsított cégében, amelynek terjeszkedése határtalannak tűnik. Én az átláthatóság, az *urbi et orbi* híve vagyok.

2019 márciusában értesítettek, hogy elnyertem az Association for Computing Machinery 2018-as évi Turing-díját, amely a számítástechnikában a Nobel-díj megfelelője. A díjon két másik, mélytanulással foglalkozó specialistával, Yoshua Bengióval és Geoffrey Hintonnal[1] osztozom, akik néha közeli, néha távolabbi kutatótársaim, de a párbeszéd soha nem szűnt meg köztünk.

A pályafutásom sokat köszönhet ezeknek a találkozásoknak, és annak a helynek, ahová fokozatosan bekerültem, az 1950-es évek zseniális kibernetikusainak örökösei közé, akik már akkoriban is olyan extravagáns és mélyreható kérdéseket vetettek fel, mint például: „Hogy lehet az, hogy a neuronok, ezek a nagyon egyszerű tárgyak, egymással összekapcsolódva létrehoznak egy olyan emergens tulajdonságot, mint az intelligencia?"

Ez a tudományos kaland tehát a lényegi kérdéseket feszegeti. Vajon egy gép, amely felismer egy autót azáltal, hogy olyan jellemzőket szűr ki, mint a kerék, a szélvédő stb., másképp működik, mint a látókérgünk, amikor azonosítja az autót? Mit kezdjünk a gép és az emberi vagy állati agy működése közötti hasonlóságokkal? A kutatási terület korlátlan.

Ám ezek a gépek, bármilyen hatékonyak és kifinomultak is, ma még rendkívül specializáltak. Összehasonlíthatatlanul kevésbé hatékonyan tanulnak, mint mi, emberek, vagy az állatok. Mind a mai napig nem rendelkeznek sem józan paraszti ésszel, sem tudattal. Legalábbis egyelőre. Kétségtelen, hogy bizonyos feladatokban felülmúlják az embereket. Már legyőzték őket go- és sakkjátékban; több száz nyelvről fordítanak; felismerik a növényeket vagy rovarokat; orvosi felvételeken kiszúrják a daganatokat. Az emberi agy azonban jelentős előnyt tudhat a magáénak: sokkalta univerzálisabb és sokkalta képlékenyebb.

Vajon mikor ugorják meg a gépek ezt a lépcsőfokot?

1
AZ MI-FORRADALOM

A mesterséges intelligencia az életünk minden területét gyarmatosítja, a gazdaságtól a kommunikáción és az egészségügyön át az önvezető járművekkel történő közlekedésig… A folyamatok láttán sokan már nem is technológiai fejlődésről, hanem egyenesen forradalomról beszélnek.

AZ MI MINDENÜTT JELEN VAN

„Alexa, milyen az idő Buenos Airesben?" Egy másodperc sem kell hozzá, és az okos hangszóró a felvett kérdést az otthoni wifin keresztül továbbítja az Amazon szervereire, azok átírják szöveggé, értelmezik, lehívják az információt a meterológiai szolgálattól, majd visszaküldik a választ, amit Alexa lágy hangon duruzsol a fülünkbe: „Jelenleg Buenos Airesben, Argentínában, 22 °C van, az égbolt felhős."

Az irodában az MI egy szorgos titkárnő. Gyorsan dolgozik, és a repetitív munka nem szegi kedvét. Több millió oldalt képes átfutni egy idézet után kutatva, és a másodperc töredéke alatt megtalálja a keresett sort, a számítógépek elképesztően gyors számítási kapacitásának köszönhetően.

Az egyik első programozható elektronikus számítógép, az ENIAC, amelyet 1945-ben építettek a Pennsylvaniai Egyetemen, hogy a lövedékek röppályáját számítsák ki vele, másodpercenként kb. 360 szorzást végzett tízjegyű számokkal. Ma már ez egy őskövületnek tűnik a személyi számítógépeink mellett, amelyek processzorai milliárdszor gyorsabbak! Több száz GFLOPS-t *(giga floating point operations per second)* képesek végrehajtani. A GPU-k *(graphical processing unit),*

amelyeket a számítógépek a grafikai megjelenítéshez használnak, néhány tucat TFLOPS *(tera floating point operations per second)* teljesítményűek.[2] Gigantikus mértékegységek, bájos kis nevekkel.

És megállíthatatlan a fejlődés! A legújabb szuperszámítógépekben több tízezer ilyen GPU van beépítve, és a kapacitásuk eléri a több százezer TFLOPS-t. Ezzel a teljesítménnyel hatalmas számítási szekvenciákat futtatnak le különféle szimulációkhoz: időjárás-előrejelzéshez, éghajlat modellezéséhez, egy repülőgép körüli légáramlás vagy egy fehérje térbeli felépítésének kiszámításához, valamint az olyan elképesztő események szimulálásához, mint amilyen az univerzum első másodpercei, a csillagok halála, a galaxisok fejlődése, az elemi részecskék ütközése vagy az atomrobbanás.

Ezek a szimulációk differenciálegyenletek vagy parciális differenciálegyenletek numerikus megoldásait számolják ki – ez a feladat korábban a matematikusokra hárult. De vajon az új számolóbajnokok ugyanolyan intelligensek, mint a hajdani matematikusok? Természetesen nem… Vagy legalábbis, még nem. A mesterséges intelligencia előtt álló egyik kihívás az, hogy egy nap képesek legyünk ezt a számolókapacitást olyan intelligens feladatok szolgálatába állítani, amelyeket általában az állatoknak és az embereknek tulajdonítunk.

A látszat persze csalóka. A mesterségesintelligencia-programok jó tanulók… egy bizonyos pontig. 2017-ben Noriko Arai a Tokiói Egyetemről, aki az informatika társadalomra gyakorolt hatásával foglalkozik, bemutatott egy MI-programot, amely átment az egyetem felvételi vizsgáján. A Todai néven futó program (Todai az egyetem beceneve) ráadásul a jelentkezők 80%-ánál jobban teljesített az eszszéírás-, a matematika- és az angolteszteken. Miközben a rendszer ostoba volt: egy szót sem értett abból, amit írt! Sikere legalább annyit elmond a japán felsőoktatási felvételi tesztek felszínességéről, mint a gépi intelligenciáról. Így aztán érthetően örömmel fogadtuk a hírt, hogy Todai végül nem nyert felvételt az egyetemre…

AZ MI MINT MŰVÉSZ

A mesterséges intelligencia egyfajta művész: másolóművész. Mesteri módon készít műveket emez vagy amaz alkotó „nyomdokain". Tetszőleges fényképet átalakít Monet-festménnyé, téli tájat átvarázsol

tavaszi jelenetté,[3] vagy egy videóban lovat kicserél zebrára. Vigyázzunk a hamisított videókkal… Mi több, 2017-ben Ahmed Elgammal csapata a Rutgers Egyetemen (USA) egy olyan rendszert képzett ki, amely teljesen új, eredeti festményeket készít, olyan jól, hogy még a szakértőket is megtéveszti.[4]

A zene sem kivétel. Számos kutató alkalmaz mesterségesintelligencia-technikákat hangszintetizálásra és zenei kompozíció előállítására. A Google Magenta projektje[5] a 2019. március 21-én bemutatott, Johann Sebastian Bach születésnapja előtt tisztelgő Google Doodle[6] révén vált ismertté: ezzel a programmal bárki a híres zeneszerző stílusában komponálhat és harmonizálhat egy dallamot. Vagy ott van az a londoni Cadogan Hallban, 2019. február 4-én adott koncert, amelyen az English Session Orchestra 66 zenésze Schubert *8. „Befejezetlen" Szimfóniája*nak egy egészen különleges változatát adta elő: azon az estén a szimfónia befejezést kapott. A két megírt tétel elemzése nyomán a Huawei-laborok MI-rendszere előállította a hiányzó két utolsó tétel dallamait. A zenekari partitúra megírásához ugyanakkor még szükség volt Lucas Cantor zeneszerző munkájára.

HUMANOIDOK? UGYAN MÁR!

Sophia, a gyönyörű kopasz nő, titokzatos mosollyal és üvegszemekkel, a 2017-es év színpadi sztárja. Életteli arcán tucatnyi váltakozó arckifejezéssel, a vicceivel – „Túl sok hollywoodi filmet nézel!", ironizál egy újságírónak, aki a robotoktól ellepett Föld miatt aggódik – annyira zavarba ejtően emberi, hogy abban az évben Szaúd-Arábia szaúdi állampolgárságot adományoz neki. Valójában azonban csak egy bábu, akinek a mérnökök tipikus válaszelemeket adtak a szájába. Amikor megszólítjuk a robotot, a beszédünket feldolgozza egy párosítórendszer, és a lehetséges válaszok katalógusából kiválasztja a legmegfelelőbb válaszokat.

Sophia becsapja a közönségét. Nem eszesebb, mint a tokiói Todai. Csak vonzóbb nála a plasztikus megjelenésével, és mi, emberek, akikre ez az animált tárgy hatással van, azt hisszük, hogy valamiféle intelligenciával rendelkezik.

KEZDETBEN VOLT A GOFAI…

A mesterséges intelligencia világa változóban van. Határai folyamatosan kitolódnak. Amikor egy probléma megoldódik, kikerül a mesterséges intelligencia területéről, és apránként beépül a klasszikus eszköztárba.

A matematikai képletek számítógéppel végrehajtható utasítássá alakítása például a mesterséges intelligencia része volt az 1950-es években, amikor a számítástechnika még gyerekcipőben járt. Ma ez a fordítóprogramok egyszerű funkciója – olyan szoftvereké, amelyek a mérnökök által írt programokat a gép által közvetlenül végrehajtható utasítássorozatokká alakítják át. Ezt az átírást manapság minden informatikushallgatónak tanítják.

Vagy vegyük az útvonalkeresést. Az 1960-as években ez egyértelműen a mesterséges intelligenciához tartozott, ma viszont már egy klasszikus modul. Léteznek hatékony algoritmusok a legrövidebb útvonal megtalálására egy gráfban (azaz utakkal összekötött pontok hálózatában), mint például az 1959-es Dijkstra-algoritmus,[7] vagy Hart, Nilsson és Raphael 1969-es A* („A star") algoritmusa.[8] A GPS-ünkkel összekapcsolva ebben a feladatban már nincsen semmi rendkívüli.

Az 1970-es és 1980-as években a mesterséges intelligencia magját a logikán és szimbólum-manipuláción alapuló automatikus következtetési technikák összessége alkotta. E hagyomány angolszász hívei némi öniróniával GOFAI-nak nevezik ezt a gyűjteményt, ami a *good old-fashioned artificial intelligence* („a jó öreg MI") rövidítése…

Ilyenek például a szakértői rendszerek: egy következtetőmotor szabályokat alkalmaz a tényekre, és ezekből új tényeket vezet le. A MYCIN-t például 1975-ben arra tervezték, hogy segítsen az orvosoknak az akut fertőzések, például az agyhártyagyulladás azonosításában, és szükség esetén antibiotikumos kezelést javasoljon. Körülbelül 600 szabályt tartalmazott, nagyjából ilyen típusúakat: „HA a fertőző organizmus Gram-negatív ÉS az organizmus pálcika alakú ÉS az organizmus anaerob, AKKOR az organizmus bakteroid (60%-os valószínűséggel)."

A MYCIN innovatív volt. A szabályai olyan bizonyosságfaktorokat tartalmaztak, amelyeket a rendszer összekombinált, hogy az eredményre vonatkozóan egy megbízhatósági pontszámot adjon. A rendszerben volt egy *„backward chaining"* („visszaláncoló") követ-

keztetési motor, amely fölállított egy – vagy több – diagnosztikai hipotézist, és kikérdezte a kezelőorvost a beteg tüneteiről. A válaszok nyomán a rendszer módosította a hipotézisét, felállította a diagnózist, majd egy bizonyos szintű megbízhatósággal javaslatot tett az antibiotikumra és az adagolásra.

Egy ilyen rendszer létrehozásához egy tudásmérnöknek le kellett ülnie az orvossal – a szakértővel – beszélgetni, és meg kellett kérnie őt, hogy részletesen magyarázza el a gondolatmenetét: hogyan diagnosztizálja a vakbélgyulladást vagy az agyhártyagyulladást? Mik a tünetek? Ebből alakultak ki a szabályok. Ha a beteg ilyen és ilyen tüneteket produkál, akkor ilyen és ilyen valószínűséggel vakbélgyulladása van, ilyen és ilyen valószínűséggel bélelzáródása van, ilyen és ilyen valószínűséggel vesebaja van. A mérnök saját kezűleg írta be ezeket a szabályokat a tudásbázisba.

A MYCIN és a későbbi rendszerek megbízhatósága nagyon jó volt, de soha nem jutottak túl a kísérleti stádiumon. A komputerizáció az orvostudományban épp csak elkezdődött, és az adatbevitel fárasztó munka volt. Még napjainkban is az. Végeredményben ezek a logikán és az elágazásos keresésen alapuló szakértői rendszerek túlontúl nehézkesnek és bonyolultnak bizonyultak. Ma már nem használják őket, de továbbra is referenciaként szolgálnak, és a mai napig szerepelnek az MI-tankönyvekben.

A logikával kapcsolatos munka mindazonáltal számos kulcsfontosságú alkalmazáshoz vezetett: a szimbolikus egyenletmegoldáshoz és integrálszámításhoz a matematikában, valamint az automatikus programverifikációhoz. Az Airbus például így ellenőrzi a repülőgépek vezérlőszoftvereinek pontosságát és megbízhatóságát.

Az MI-kutatók egy része továbbra is ezeken a témákon dolgozik, míg egy másik részük, akikhez én is tartozom, a gépi tanuláson alapuló, nagyon eltérő megközelítésekkel foglalkozik.

...MAJD JÖTT A GÉPI TANULÁS

Az emberi intelligenciának csak egy kis részét teszi ki az okfejtés és a racionális következtetés. Gyakran analógiák alapján gondolkodunk, intuitív módon cselekszünk, a világról alkotott reprezentációkra támaszkodva, amelyeket a tapasztalatok révén apránként sajátítottunk

el. Észlelés, intuíció, tapasztalat: ezek mind-mind tanult képességek, vagy legalábbis a gyakorlatban csiszolódó képességek.

Ilyen körülmények között, ha olyan gépet akarunk építeni, melynek intelligenciája megközelíti az emberi intelligenciát, akkor alkalmassá kell tennünk arra, hogy tanuljon. Az emberi agy kb. 86 milliárd egymással összekapcsolt neuron (azaz idegsejt) hálózatából áll, és ebből 16 milliárd az agykéregben található. Minden egyes neuron átlagosan közel 2000 másik neuronhoz kapcsolódik a szinapszisoknak nevezett kapcsolatokon keresztül. Az agyban a tanulás szinapszisok létesítésével, szinapszisok megszüntetésével vagy hatékonyságuk módosításával történik. A gépi tanulás manapság legdivatosabb megközelítésében mesterséges neurális hálózatokat építenek, és a tanulási folyamat során módosulnak a neuronok közötti kapcsolatok.

Íme, néhány általános elv e tanulóhálózatokkal kapcsolatban:

A *gépi tanulás*nál van egy első fázis, a tanulás avagy tréningezés szakasza, melynek során a gép fokozatosan „megtanul" teljesíteni egy feladatot, majd következik a második fázis, a működési szakasz, amikor a gép már nem tanul, csak elvégzi a feladatot.

Ahhoz, hogy egy gépet megtanítsunk arra, hogy meg tudja állapítani, hogy egy képen autó vagy repülőgép van-e, először is több ezer olyan képet kell mutatnunk neki, amelyeken repülőgép vagy autó van. Minden alkalommal, amikor a gép beolvassa a képet, az egymással összekapcsolt mesterséges neuronokból álló belső hálózata – ami a valóságban a számítógép által kiszámított matematikai függvényeket jelenti – feldolgozza az illető képet, és ad egy kimeneti választ. Ha a válasz helyes, akkor nem történik semmi, és továbblépünk a következő képre. Ha a válasz helytelen, kissé módosítjuk a gép belső paramétereit, azaz a neuronok közötti kapcsolatok erősségét, hogy a gép kimenete közelebb kerüljön az elvárt válaszhoz. Idővel a rendszer alkalmazkodik, és végül bármilyen tárgyat felismer, legyen az egy korábban látott kép, vagy bármilyen más kép. Ezt nevezzük általánosító képességnek.

Ezt a folyamatot, hamarosan látni fogjuk, az agy működése ihlette, de a gépek még fényévekre le vannak maradva az agytól. Csak néhány számadat: az agy 86×10^9 neuronból áll, amelyeket körülbelül $1{,}5 \times 10^{14}$ szinapszis köt össze egymással. Minden egyes szinapszis másodpercenként akár százszor is képes „számítást" végezni. Ez a szinaptikus számítás kb. száz numerikus számítási műveletnek

(szorzásnak, összeadásnak stb.) felel meg egy számítógépben, ami másodpercenként $1,5 \times 10^{18}$ műveletet jelent az egész agyban. A valóságban a neuronoknak mindig csak egy része aktiválódik. Összehasonlításképpen: egy GPU-kártya másodpercenként 10^{13} műveletet képes végrehajtani. Vagyis 100 000 GPU-kártyára lenne szükségünk ahhoz, hogy megközelítsük az agy teljesítményét! Ám van egy bökkenő: az emberi agy 25 wattnak megfelelő energiát fogyaszt, míg egyetlen GPU-kártya ennek a tízszeresét, 250 wattot fogyaszt! Az elektronika nagyságrendekkel kevésbé hatékony, mint a biológia.

A RÉGI ÉS A MODERN KOKTÉLJA

A mai alkalmazások általában a gépi tanulás, a GOFAI és a klasszikus számítástechnika keverékei. Gondoljunk csak az önvezető autóra: a fedélzeti vizuális felismerőrendszer, amely az úton lévő tárgyak és vizuális jelzések észlelésére, lokalizálására és beazonosítására van kiképezve, egy különleges neurális hálózati architektúrát, az úgynevezett „konvolúciós hálózatot" használja, míg a döntés, amelyet az autó hoz, miután „meglátta" a sávjelzéseket, a járdát, a parkoló autót vagy a kerékpárt, a hagyományos, kézzel írt pályatervező rendszerektől, vagy akár GOFAI-nak tekinthető szabályalapú rendszerektől függ.

Ezek a teljesen autonóm járművek még tesztelés alatt állnak, de a kereskedelmi forgalomban kapható autók, mint például a Tesla, 2015 óta már rendelkeznek olyan vezetéstámogató rendszerekkel, amelyek konvolúciós hálózatokat használnak. A látórendszerekkel felszerelt sebességszabályozó automaták a járművet az autópályán önvezető üzemmódba kapcsolják, sávon belül tartják, és automatikusan kihúzódnak a sávból, ha a vezető bekapcsolja az irányjelző villogót, és közben azt is érzékelik, hogy jön-e az autó mögött másik jármű.

EGY DEFINÍCIÓS KÍSÉRLET

A könyvben végig ezen a területen fogunk kalandozni, ám most itt az ideje, hogy egy pillanatra megálljunk, és kissé hátrébb lépjünk. Hogyan határozhatjuk meg ezeknek a mesterségesintelligencia-rendszereknek a közös jellemzőit?

Én azt mondanám, hogy a mesterséges intelligencia egy gép azon képessége, hogy olyan feladatokat végezzen, amilyeneket általában az állatok és az emberek végeznek: észleljen, gondolkodjon és cselekedjen. Ez elválaszthatatlan az élőlényeknél megfigyelhető tanulási képességtől. A mesterségesintelligencia-rendszerek nem mások, mint rendkívül kifinomult elektronikus áramkörök és számítógépes programok. Tárolási és memóriaelérési kapacitásuk, számítási sebességük és tanulási képességük azonban lehetővé teszi számukra, hogy „absztrahálják" a hatalmas adatmennyiségekben foglalt információkat.

Észlelni, gondolkodni és cselekedni. Alan Turing, az angol matematikus, aki a számítástechnika úttörője volt, és a második világháború idején megfejtette a német hadsereg Enigma-kódrendszerét, valóságos látnok volt. Már akkoriban megérezte a tanulás fontosságát, amikor ezt írta: „Ahelyett, hogy megpróbálnánk egy olyan programot készíteni, amely a felnőtt elmét szimulálja, miért ne próbálnánk meg egy olyat készíteni, amely a gyermek elméjét szimulálja? Ha ezt azután megfelelő oktatásnak vetnénk alá, megkapnánk a felnőtt agyat."[9]

Alan Turing nevéhez fűződik a híres teszt is, amely egy ember és két, általa nem látott beszélgetőpartner – egy számítógép és egy másik ember – közötti írásos párbeszédből áll.[10] Ha meghatározott idő elteltével a tesztelő személy nem veszi észre, hogy a két fél közül melyik a gép, akkor a gép átment a teszten. Ma azonban a mesterséges intelligencia olyannyira fejlett, hogy ezt a tesztet a kutatók már nem tartják relevánsnak. A társalgási képesség az intelligenciának csak egy nagyon speciális formája. Egy MI-rendszer pedig könnyen azt a benyomást keltheti, hogy intelligens. Csak annyit kell tennie, hogy egy kissé szétszórt, kissé autista, kelet-európai tinédzsernek adja ki magát, aki nem beszél túl jól angolul, és a hallgatóság máris megbocsátja neki a félreértéseket és a szintaktikai baklövéseket…

A JÖVŐBELI FEJLŐDÉS IRÁNYA

Meggyőződésem, hogy a *deep learning* a mesterséges intelligencia jövőjének része. Ma azonban egy mélytanuló rendszer nem képes arra, hogy logikusan gondolkodjon, míg a logika a jelenlegi formájában nem kompatibilis a tanulással. Az elkövetkező évek kihívása, hogy a logikát és a mélytanulást kompatibilissé tegyük egymással.

A mélytanulás mindenesetre nagyon hathatós eszköz… de egyszersmind nagyon korlátozott is. Kizárt például, hogy egy sakkozásra betanított gép gojátékot játsszon, vagy fordítva. A gép úgy lép, hogy a leghalványabb fogalma sincs arról, mit csinál, és jelenleg kevesebb a sütnivalója, mint egy kóbor cicának. Ha az MI-rendszereket az értelmi képességek legendás mérőskáláján kellene elhelyeznünk, ahol az ember a 100-as értéknél, az egér pedig az 1-esnél van, akkor közelebb lennének a kis rágcsálóhoz. És ez még akkor is így van, ha a mesterséges intelligencia teljesítménye a precíz, behatárolt feladatokban túlszárnyalja az embert.

AZ ALGORITMUSOK NAGY ÁRIÁJA

Az algoritmus utasítások sorozata. Se több, se kevesebb. Nincs benne semmi varázslatos és semmi rejtélyes. Mondok egy példát. Van egy számlistám, amit növekvő rendbe akarok rendezni. Írok egy számítógépes programot, amely beolvassa az első számot, összehasonlítja a másodikkal, és ha az első nagyobb, mint a második, akkor felcseréli őket. Ezután összehasonlítom a második és a harmadik számot, és ugyanezt a műveletet ismétlem addig, amíg el nem érek a lista utolsó számához. Ezután újra végigmegyek a listán, annyiszor, ahányszor csak szükséges, amíg végül már egyetlen pár sem lesz fordított helyzetű.

Ezt a számlistarendező algoritmust „buborékrendezésnek” hívják. Egy fiktív programnyelven pontos utasítássorozattá tudom lefordítani.[11]

```
Buborékrendezés (T lista)
   ciklus i = (T mérete –1)-től 1-ig
            ciklus j = 0-tól (i – 1)-ig
                     ha T[j + 1] < T[j] akkor
                              csere (T,j + 1,j)
```

Vegyél egy értéket, hasonlítsd össze egy másikkal, add hozzá egy harmadikhoz, végezd el a matematikai műveleteket, hajtsd végre a ciklusokat, vizsgáld meg, hogy egy feltétel igaz-e vagy hamis stb. Az algoritmus nem más, mint egy recept.

Manapság gyakran beszélnek az emberek a Facebook „algoritmusáról" vagy a Google „algoritmusáról", ez azonban visszaélés a szavakkal. A Google keresőoldal esetében inkább „algoritmusok gyűjteményéről" van szó, amely listát készít az összes olyan weboldalról, amely tartalmazza a keresett szöveget. Több százról, vagy akár több ezerről! Minden ilyen weboldalhoz más algoritmusok – kézzel írt vagy betanított algoritmusok – által előállított pontszámok sorozata tartozik. Ezek a pontszámok az illető webhely népszerűségét, megbízhatóságát, tartalmi relevanciáját értékelik, azt, hogy ha a keresett kifejezés egy kérdés, tartalmaz-e rá választ, és azt, hogy a tartalma a felhasználó érdeklődési körének megfelelő-e. Ez meglehetősen komplex dolog.

Ha viszont csak a betanított rendszereket nézzük, a programkód, amely futtatja őket és kiszámítja a pontszámokat, elvileg pofonegyszerű, néhány sorban elférne, ha nem volna fontos szempont a futási sebesség (valójában amiatt lesz kissé bonyolult a kód, mivel azt szeretnénk, hogy gyorsan fusson). A rendszer valódi bonyolultsága a hálózat neuronjai közötti kapcsolatokban rejlik, amelyek a hálózat architektúrájától és a tanításától függnek, nem pedig a kimeneteket kiszámító programkódban.

Mielőtt rátérnék az intelligens gép belső működésének részletes vizsgálatára, szeretném felvázolni a mesterséges intelligencia történetét a XX. század közepétől. Ez egy lenyűgöző kaland, amelybe én már fiatalon belecsöppentem, telis-tele ádáz vitákkal és nagyszabású víziókkal, viharos időszakokkal és hosszú szélcsendekkel, és amelyben a gépi logika mellett hitet tevő tudósok útját keresztezik azok a kutatók, akik az idegtudományok és a kibernetika ihlette tanulási képességek fejlesztésén dolgoznak, köztük én is.

2
RÖVID ÖSSZEFOGLALÓ AZ MI TÖRTÉNETÉRŐL… ÉS A SAJÁT PÁLYAFUTÁSOMRÓL

AZ ÖRÖKÖS KERESÉS

A mesterséges intelligencia történetének eredete, mint azt Pamela McCorduck amerikai író írja, „az ember régi vágyakozása, hogy Istent játszszon". Az ember már régóta próbál olyan automatákat létrehozni, amelyek az élet illúzióját keltik. A XX. században a tudomány és az idegtudományok fejlődése még azt a reményt is felcsillantotta, hogy a gondolkodási folyamatot gépesíteni lehet: az 1950-es években az első robotok és számítógépek megjelenésével egyes utópisták azt jósolták, hogy a gép hamarosan kifejleszti majd az emberi intelligenciát. A sci-fi irodalomban konkrétan megelevenedtek ezek az álmok. Ám ma már messze vagyunk ettől.

Ebben a hosszú kalandban a fejlődés a technikai újításoktól függ: a gyorsabb számítógépektől, az egyre több adat egyre kisebb méretre zsugorodó tárolásától. 1977-ben a Cray-1 szuperszámítógép 160 MFLOPS[12] számítási kapacitással rendelkezett, 5 tonnát nyomott, 115 kilowattot fogyasztott, és 8 millió dollárba került. Ma egy 300 eurós grafikus kártya, amely számos, játékokra szánt PC-ben megtalálható, 10 TFLOPS[13] – azaz 60 000-szer nagyobb – teljesítményre képes. Hamarosan minden okostelefon hasonló teljesítményű lesz.

Ha a történetnek szüksége van egy kezdetre, hát kezdjük a dartmouth-i konferenciával, ahol a mára ikonikussá vált „mesterséges intelligencia" kifejezés megszületett. A konferenciát 1956 nyarán a Hanover (New Hampshire) melletti Dartmouth College-ban szervezte két úttörő tudós, Marvin Minsky és John McCarthy. Marvin Minskyt lenyűgözték a tanulásra képes gépek. 1951-ben egy princetoni diáktársával megépítette az egyik első neurális gépet, a SNARC-ot,

egy kezdetleges, 40 „szinapszissal” felszerelt, tanulásra képes, neurális hálózati áramkört. John McCarthy találta fel a LISP-et, a mesterséges intelligenciában széles körben használt programozási nyelvet. Az ő nevéhez fűződik a sakkprogramoknál alkalmazott fabejárási algoritmus is (*tree traversal algorithm* – szoktak fakeresési algoritmusként is hivatkozni rá). Mintegy húsz kutató válaszolt a konferenciafelhívásra, köztük Claude Shannon, a Bell Labs (az AT&T óriás telefontársaság New Jersey-i laboratóriumai) villamosmérnöke és matematikusa, Nathan Rochester az IBM-től és Ray Solomonoff, a gépi tanulás elméletének úttörője. A konferencián eszmecserét folytattak az informatika és a kibernetika virágzó területeiről: a természetes és mesterséges rendszerek szabályozásának tanulmányozásáról, a komplex információfeldolgozásról, a mesterséges neurális hálózatokról, az automaták elméletéről és egyéb témákról. Ez a miniatűr szimpózium megfogalmazta azokat az alapelveket, amelyek a John McCarthy által „mesterséges intelligenciának” nevezett tudományterület születési anyakönyvét jelentették.

ELSŐSORBAN LOGIKA

Abban az időben egyes tudósok kizárólag logikai alapú intelligens gépeket tudtak elképzelni: fabejáráson és szakértői rendszereken alapuló gépeket. A mérnök igaz tényeket és szabályokat ad meg nekik, a gépek pedig ezekből következtetnek más tényekre. A cél egy olyan gép megépítése, amely az embert helyettesíti az összetett gondolkodásban. Allan Newell és Herbert Simon, a pittsburghi Carnegie Mellon Egyetem munkatársai mutatták meg az utat Logic Theorist nevű programjukkal, amely egyszerű matematikai tételeket tudott bizonyítani matematikai képletek átalakításaiból álló fák feltárásával. Ez volt a nagy elvárások időszaka.

Aztán a mesterséges intelligenciára beköszöntött az első „tél”. Az amerikai védelmi minisztérium ARPA[14] ügyosztálya 1970-ben csökkentette a mesterséges intelligencia kutatására szánt költségvetést. Három évvel később az Egyesült Királyság is így tett a Lighthill-jelentést követően, amely lehűtötte a tudományos lelkesedést. És ugyebár több pénz több kutatást jelent…

Az 1980-as évek elején fordultak a dolgok. A szakértői rendszerek komoly reményeket keltettek, és Japán elindította a nagyra törő „ötödik generációs" számítógép-projektjét, amelynek már a konstrukciójába is logikai következtetési képességeket kellett volna integrálni. Képesnek kellett volna lennie arra, hogy beszélgessen, szöveget fordítson, képeket értelmezzen, és talán még arra is, hogy úgy gondolkodjon, mint egy ember. Sajnos azonban ez a projekt fiaskó lett. A korábban már említett MYCIN és a hozzá hasonló szakértői rendszerek kifejlesztése és forgalmazása a vártnál nehezebbnek bizonyult. Nem működik jól az az elv, hogy „tudásmérnökök" ülnek az orvosok vagy konstruktőrök mellett, és megkísérlik rögzíteni, hogy milyen gondolatmenetet követnek, amikor egy betegséget próbálnak azonosítani vagy egy hibát próbálnak diagnosztizálni. A vártnál bonyolultabb, drágább és kevésbé megbízhatóbb a szakember által mozgósított összes tudást egy szabályrendszerre redukálni.

E nehezen reprodukálható klasszikus intelligencia mellett meg kell még említenünk a fabejárási módszert, amely átütő sikereket ért el.

A JÁTÉKOK VILÁGA

1997-ben Garri Kaszparov sakkvilágbajnok New Yorkban hatjátszmás visszavágót játszott a multinacionális IBM vállalat által programozott szuperszámítógép, a Deep Blue ellen. A gép egy majd' 2 méter magas, 1,4 tonnás monstrum volt. A visszavágó hatodik játszmájában, három döntetlen játszmát követően, Garri Kaszparov mindössze 19 lépés után abbahagyta a játékot, tehetetlenül szétttárta karját, ami a sakkrajongók millióiban hagyott mély nyomot. Elismerte, hogy alulmaradt a gép helyzetfeltáró kapacitásával szemben.

Vessünk egy pillantást a Deep Blue-ra! A gép 30 processzorral, plusz 480 olyan speciális áramkörrel rendelkezett, amelyeket kifejezetten a sakkállások vizsgálatára terveztek. Ennek a számítási kapacitásnak köszönhetően a gép másodpercenként mintegy 200 millió sakkállás minőségét tudta kiértékelni lényegében a klasszikus fakeresési technikával.

Néhány évvel később, 2011. február 14-én, 15-én és 16-án az IBM Watson számítógépe nyerte meg a *Jeopardy!* című amerikai műveltségi vetélkedőt – három forduló után. A két bajnok között elhelyezkedő

számítógép avatárját egy fénysugarakkal tarkított földgömb jelenítette meg. Az informatikusok által írt program a feltett kérdésből kiszűrte a felesleges kifejezéseket (cikkek, prepozíciók stb.), azonosította a lényeges szavakat, majd a helyes válasz kiválasztásához hatalmas mennyiségű, 200 millió oldalnyi szövegben kereste meg ezeket a szavakat, beazonosítva azokat a mondatokat, amelyekben a válasz megtalálható. Ezeket a szövegeket – a teljes Wikipédiát, enciklopédiákat, szótárakat, lexikonokat, sajtóközleményeket és irodalmi műveket – a 16 terabájtos[15] RAM-memóriában tárolta (a merevlemezek túl lassúak voltak ehhez az alkalmazáshoz). A több százezer tárgyszó és tulajdonnév egy nagy listát alkotott, amelyek mindegyike egy adott Wikipédia-cikkre, weboldalra vagy szövegre utalt, amelyben megjelent… A Watson rendszer ellenőrizte, hogy van-e olyan dokumentum, amelyben előfordul az adott kérdésben szereplő kulcskifejezések valamelyike. A feladat ezután az volt, hogy megkeresse a megfelelő választ a cikkben.

Ha például a kérdés az volt, hogy „Hol született Barack Obama?", Watson tudta, hogy a válasz egy helységnév lesz. Az adatbázisában volt egy lista az összes olyan dokumentumról, amely Barack Obamát említette. Tehát valahol biztosan van olyan cikk, amely az „Obama", „született" és „Hawaii" szavakat tartalmazza. A gépnek csak annyit kellett tennie, hogy kiválasztja a „Hawaii" szót, ami a helyes válasz. Watson lényegében egy gyors információkereső rendszer volt, jó adatindexálással párosítva. De ez a rendszer nem értette a kérdés jelentését. Úgy viselkedett, mint egy kisiskolás gyerek, aki a tankönyvét lapozgatva (vagy épp a Wikipédiát böngészve) írja meg a házi feladatát. Egy adott kérdésre képes megtalálni a helyes választ a tankönyvben, képes kimásolni azt, ám vajmi keveset ért abból, amit leír.

Egy újabb bravúr jött 2016-ban. Szöulban a dél-koreai gobajnok kikapott számítógépes ellenfelétől, az AlphaGótól, a Google leányvállalata, a DeepMind által tervezett impozáns rendszertől. Lee Sedol, a tizennyolcszoros világbajnok, ötből négy játszmát veszített el a programmal szemben. A Deep Blue-val ellentétben az AlphaGo „tanuláson" ment át. A gép úgy képezte ki magát gojátékossá, hogy saját maga ellen játszott, és több már jól ismert gépi technikát integrált: a konvolúciós hálózatokat, a megerősítéses tanulást és a Monte-Carlo fakeresést, amely egy véletlenszerű fabejárási módszer. De ne szaladjunk ennyire előre…

AZ IDEGTUDOMÁNY ÉS A PERCEPTRON

Az 1950-es években, miközben a klasszikus, logikán és fabejáráson alapuló mesterséges intelligencia hírnökei feszegették annak határait, a tanulás úttörői kezdték hallatni a hangjukat. Megvédték azt a gondolatot, hogy ha azt akarjuk, hogy a számítógépes rendszerek az állatokhoz és az emberekhez hasonlóan komplex feladatokra legyenek képesek, akkor a logika nem elég. Közelebb kell kerülnünk az agy működéséhez, és ezért olyan rendszereket kell létrehoznunk, amelyek képesek önmagukat programozni, inspirálódva annak tanulási mechanizmusaiból. Én magam is ezen a mélytanuláson és (mesterséges) neurális hálókon alapuló kutatási területen dolgozom. Ez működik az összes mai látványos alkalmazásban, így például az önvezető autóban is.

Ennek a kutatási csapásnak az eredete a XX. század közepére nyúlik vissza. Az utópisták a mesterséges intelligencia területén már az 1950-es években magukévá tették Donald Hebb kanadai pszichológus és neurobiológus elméleteit, aki a neuronális kapcsolatok tanulásban betöltött szerepén töprengett. Ahelyett, hogy az emberi gondolkodás logikai szekvenciáit reprodukálnánk, miért ne tárnánk fel inkább azok szubsztrátumát, azt a nagyszerű biológiai processzort, ami az agy?

A neurális áramlathoz tartozó informatikusok (szemben a korábbi logikai vagy „szekvenciális" áramlattal) ezért a biológiai idegkörök modellezésére törekednek. A gépi tanulás, amelyre erőfeszítéseik irányulnak, egy eredeti architektúrán, matematikai függvények hálózatán alapul, amelyet analógiával „mesterséges neuronoknak" neveznek. A hálózat megkapja a bemeneti jelet, és ezt a mesterséges neuronok úgy dolgozzák fel, hogy a kimeneten a jel beazonosításra kerül. A komplexitás – egy alakzat felismerése – nagyon egyszerű elemeknek, a mesterséges neuronoknak az együttes működéséből származik. Ugyanúgy, mint az agyban, ahol az alapvető funkcionális egységek, a neuronok kölcsönhatása hozza létre a gondolkodást.

Ennek az irányzatnak az alapműve 1957-re nyúlik vissza: abban az évben a Cornell Egyetemen Frank Rosenblatt pszichológus megépítette a perceptront, az első tanulógépet, amelyet Donald Hebb kognitív elmélete ihletett. A következő fejezetben foglalkozunk majd vele, mivel a mai napig a gépi tanulás egyik etalonja. A perceptron egy betanítási időszak után képes formákat felismerni.

Az 1970-es években két amerikai, Richard Duda, aki akkoriban a San José-i Egyetem (Kalifornia) villamosmérnöki professzora volt, és Peter Hart, az SRI (Stanford Research Institute) informatikusa a Menlo Parkban (Kalifornia) számba vette az úgynevezett „statisztikus mintafelismerési" módszereket, amelyek közül a perceptron csak egy példa.[16] Tankönyvük azonnal referenciává vált a mintafelismerés világában, bibliája lett a diákoknak… És nekem is.

De a perceptron nem varázsszer. Az egyetlen mesterséges neuronrétegből álló rendszer korlátozott képességekkel rendelkezik. A kutatók úgy próbálták javítani a hatékonyságát, hogy egy helyett több neuronréteget illesztettek bele. De nem sikerült megtalálniuk azt az algoritmust (a híres utasítássorozatot), amely lehetővé tette volna, hogy az összes réteget betanítsák… így a gép továbbra is nagyon korlátozott maradt.

BEKÖSZÖNT AZ ÁLTALÁNOS TÉL

Még mindig ebben az időszakban járunk, amikor 1969-ben Seymour Papert és Marvin Minsky – ugyanaz a tudós, aki az 1950-es években lelkesedett a mesterséges neurális hálózatokért, majd később elfordult tőlük – megjelentette a *Perceptrons: An Introduction to Computational Geometry* című könyvet.[17] Ebben bemutatták a tanulógépek korlátait, amelyek közül néhány lényegi és orvosolhatatlan. Véleményük szerint az egész kaland halálra van ítélve. Ez a két MIT-professzor nagy tekintélynek örvendett, és a könyvük nagy port kavart. A kutatásfinanszírozó ügynökségek leállították az ezen a területen folytatott munkák támogatását. A GOFAI-kutatáshoz hasonlóan a neurális hálózatok kutatásánál is beköszöntött az első „tél".

A legtöbb tudós ekkoriban már nem arról beszélt, hogy intelligens, tanulásra képes gépeket akar építeni, hanem inkább földhözragadtabb projektekre korlátozták az ambícióikat. A neurális hálózatoktól örökölt módszereket használva létrehozták például az „adaptív szűrést", a modern világ számos kommunikációs technológiája mögött rejlő folyamatot. Korábban, amikor azt akartuk, hogy két számítógép telefonvonalon keresztül cseréljen adatokat, a vonal tulajdonságai olyanok voltak, hogy ha a bemeneten egy bináris jelet küldtünk, amelynek a feszültsége 0-tól 48 voltig terjedt, akkor az leromlott, mire elérte

a néhány kilométerre lévő célállomást. Manapság adaptív szűrővel történik a helyreállítása. Az alkalmazott algoritmust Lucky-algoritmusnak hívják, a feltalálójáról, Bob Luckyról nevezték el, aki mintegy 300 fős részleget vezetett a Bell Labsben, ahol az 1980-as évek végén magam is dolgoztam.

E nélkül az adaptív szűrés nélkül nem lenne hangszóróval felszerelt telefonunk, amely lehetővé teszi, hogy a mikrofonba beszéljünk anélkül, hogy a mikrofon rögzítené a beszélgetőpartnerünk hangját is (ami néha előfordul: halljuk magunkat beszélni). A visszhangszűrők a perceptron algoritmushoz nagyon hasonló algoritmusokat használnak. És nem lenne modemünk sem.[18] A modem lehetővé teszi, hogy egy számítógép telefonvonalon vagy bármilyen más kommunikációs vonalon keresztül kommunikáljon egy másik számítógéppel.

HOLDKÓROSOK

Az 1970-es és 1980-as évek nagy telén néhány ember továbbra is kitartóan dolgozott a neurális hálózatokon, de a tudományos közösség csak holdkórosoknak *(„lunatic fringe")* nevezte őket. A finn Teuvo Kohonenre gondolok, aki a neurális hálózatokhoz szorosan kapcsolódó „asszociatív memóriákról" írt, valamint egy sereg japán kutatóra – a japánok elszigetelt mérnöki tudományos ökoszisztémával rendelkeztek, amely távol áll a nyugatiakétól –, köztük Shun-Ichi Amari matematikusra és egy bizonyos Kunihiko Fukushimára. Az utóbbi egy olyan gépen dolgozott, amelyet Cognitronnak nevezett, a perceptron kifejezéssel játszva. Két változatot készített: az 1970-es évekbeli Cognitront és az 1980-as évekbeli Neocognitront. Akárcsak annak idején Rosenblattot, Fukushimát is az idegtudományok fejlődése inspirálta, különösen az amerikai David H. Hubel és a svéd Torsten N. Wiesel munkássága.

Ez utóbbi két neurobiológus 1981-ben élettani Nobel-díjat kapott a macska látórendszerével kapcsolatos munkájáért. Kimutatták, hogy a látás úgy jön létre, hogy a vizuális jel több neuronrétegen halad keresztül: a retinából átkerül az elsődleges látókéregbe, onnan a látókéreg más területeire, majd végül eljut az inferotemporális kéregbe. Ezekben a rétegekben ráadásul a neuronoknak nagyon specifikus funkcióik vannak. Az elsődleges látókéregben minden egyes neuron

a látómezőnek csak egy kis területéhez, a receptív mezőjéhez kapcsolódik. Ezeket a neuronokat egyszerű sejteknek nevezzük. A következő rétegben más egységek integrálják az előző réteg aktivációit, ami lehetővé teszi a kép reprezentációjának megőrzését, ha a tárgy kissé elmozdul a látómezőben. Ezeket az egységeket komplex sejteknek nevezzük.

Fukushimát tehát ez az elképzelés inspirálta, miszerint az első réteg egyszerű sejtekből áll, amelyek a képet szegélyező kis receptív mezőkben egyszerű mintákat érzékelnek, a következő rétegben pedig komplex sejtek találhatók. A Neocognitron összesen öt rétegből állt: egyszerű sejtekből, komplex sejtekből, egyszerű sejtekből, komplex sejtekből, majd egy perceptronhoz hasonló osztályozó rétegből. Az első négy réteghez egyfajta tanulási algoritmust használt, de ez az algoritmus „felügyelet nélküli" volt, ami azt jelenti, hogy nem vette figyelembe az elvégzendő végső feladatot. Ezeket a rétegeket „vakon" képezték. Csak az utolsó réteget képezték felügyelt módon, mint a perceptron esetében. Fukushima nem rendelkezett olyan tanulási algoritmussal, amely lehetővé tette volna, hogy a Neocognitron összes rétegének paramétereit beállítsa. Hálózata azonban képes volt bizonyos egyszerű alakzatok, például számok felismerésére.

Az 1980-as évek elején Fukushima nem volt teljesen egyedül ezen a kutatási területen. Számos észak-amerikai csapat is dolgozott rajta: pszichológusok, mint Jay McClelland és David Rumelhart, biofizikusok, mint John Hopfield és Terry Sejnowski, és informatikusok, mint Geoffrey Hinton. Utóbbival osztoztam 2018-ban a Turing-díjon.

SZÍNRE LÉPEK

Jómagam az 1970-es években kezdtem el érdeklődni ezen témák iránt. Talán az keltette fel a kíváncsiságomat, hogy figyeltem édesapámat, aki repülőmérnökként és zseniális barkácsolóként a szabadidejében elektronikával foglalkozott. Távirányítású repülőmodelleket épített. Emlékszem, az első távirányítóját egy kisautó és egy csónak irányítására az 1968-as májusi sztrájkok idején készítette, mert otthon ragadt. Nem én vagyok az egyetlen a családban, akinek továbbadta a lángot. Az öcsém, aki hat évvel fiatalabb nálam, szintén informatikus lett, ő az akadémiai karrier után most a Google kutatója.

Már nagyon korán lenyűgözött a technológia, az űr meghódítása és a számítástechnika kezdetei. Arról álmodtam, hogy paleontológus leszek, mert lenyűgözött az emberi intelligencia kialakulása és evolúciója. Még ma is úgy gondolom, hogy agyunk teljesítménye a legrejtélyesebb dolog az élővilágban. Emlékszem, hogy a *2001: Űrodüsszeia* című filmet moziban láttam Párizsban, a szüleimmel, a nagynénémmel és a nagybátyámmal, akik sci-fi rajongók voltak. Nyolcéves voltam akkoriban. A film mindent érintett, amit szerettem: az űrutazást, az emberiség jövőjét és a HAL nevű szuperszámítógép lázadását. Vajon HAL hajlandó-e ölni, hogy biztosítsa saját túlélését és a küldetés sikerét? Már akkor is lenyűgözött a kérdés, hogyan lehet az emberi intelligenciát egy gépben reprodukálni.

Természetesen a középiskola után gyakorlati projekteken akartam dolgozni. 1978-ban beiratkoztam egy párizsi elektronikai mérnöki iskolába, az ESIEE-be, ahová közvetlenül az érettségi után lehetett jelentkezni, előkészítő kurzusok nélkül. Vegyük tudomásul, az előkészítő kurzusok elvégzése nem az egyetlen módja annak, hogy a tudományokban sikeresek legyünk. Tanúsíthatom! És mivel az ESIEE-n folytatott tanulmányaim nagyfokú önállóságot biztosítottak számomra, ezt ki is használtam a későbbiekben!

GYÜMÖLCSÖZŐ OLVASMÁNYOK

Az egyik első dolog, ami elvarázsolt, a Cerisy konferencián a velünk születettség versus szerzettség vitáról szóló beszámoló volt, amelyet 1980-ban fedeztem fel.[19] Noam Chomsky nyelvész azt állította, hogy az agyban vannak olyan, már eleve meglévő struktúrák, amelyek jóvoltából meg tudunk tanulni beszélni. Jean Piaget fejlődéspszichológus viszont azt az álláspontot védte, hogy minden tanulható, így a struktúrák egy része is, és a nyelvet az intelligencia fejlődésével fokozatosan sajátítjuk el. Az intelligencia eszerint tehát a külvilággal való cserekapcsolatokon alapuló tanulás eredménye. Ez az elképzelés tetszett nekem, és azon tűnődtem, vajon hogyan lehetne a gépekre alkalmazni. Ebben a vitában neves tudósok vettek részt, köztük Seymour Papert, aki a szóban forgó könyvben a perceptron dicshimnuszát zengte, úgy írta le, mint egy egyszerű gépet, amely képes összetett feladatokat megtanulni.

Felfedeztem, hogy létezik ilyen tanulógép. A kérdéskör lenyűgözött! Mivel szerda délutánonként nem volt tanítás, a Rocquencourt-ban székelő INRIA (a francia Nemzeti Számítástechnikai és Irányítástechnikai Kutatóintézet) könyvtárának szakkönyveit lapozgattam. Ez az intézmény rendelkezik a leggazdagabb informatikai gyűjteménnyel a párizsi régióban. Hamarosan ráébredtem, hogy a nyugati világban már senki sem foglalkozik neurális hálózatokkal. És döbbenten vettem észre, hogy a könyv, amely a perceptronnal kapcsolatos kutatásokat leállította, ugyanettől a Seymour Paperttől származik!

A másik szenvedélyemmé a rendszerelmélet vált, amely a mesterséges rendszereket és a természetes biológiai rendszereket tanulmányozza. Ezt az 1950-es években kibernetikának nevezték. Például ott van a testhőmérséklet-szabályozó rendszer: a test egyfajta termosztátnak köszönhetően tartja 37 °C fokon a hőmérsékletét, amely korrigálja a test és a külvilág hőmérséklete közötti különbséget.

Az önszerveződés magával ragadt. Hogyan lehetséges, hogy a molekulák vagy viszonylag egyszerű tárgyak spontán módon összetett struktúrákká szerveződnek? Hogyan keletkezhet intelligencia egyszerű, egymással kölcsönhatásban álló elemek, ti. neuronok óriási halmazából?

Kolmogorov, Solomonoff és Chaitin algoritmikus komplexitáselméletével kapcsolatos matematikai munkákat tanulmányoztam. Duda és Hart könyve,[20] amelyet korábban már említettem, a legfontosabb olvasmányommá vált, mindig ott volt az éjjeliszekrényemen. Olvastam a *Biological Cybernetics* című folyóiratot, amely az agy és az élő rendszerek működésének számítógépes matematikai modelljeivel foglalkozott.

Ezek a mesterséges intelligencia téli álma idején elhanyagolt kérdések megszólítottak engem, és fokozatosan kialakult bennem a meggyőződés, hogy ha intelligens gépeket akarunk építeni, nem elég, ha logikusan működnek, arra is képesnek kell lenniük, hogy tanuljanak, vagyis a tapasztalatból kiindulva konstruálják meg magukat.

Olvasmányaim során tudatosult bennem, hogy a tudományos közösség egy része osztja ezt a nézetet. Felfedeztem Fukushima munkásságát, és azon töprengtem, hogyan lehetne javítani a Neocognitron neurális hálózatainak hatékonyságát. Szerencsére az ESIEE biztosított a hallgatók számára akkoriban igen nagy teljesítményűnek számító komputereket. Programokat írtunk egy iskolai barátommal, Philippe

Metsuval, aki hozzám hasonlóan rajongott a mesterséges intelligenciáért, de őt inkább a gyerekkori tanulás pszichológiája érdekelte. Az iskola néhány matematikatanára beleegyezett, hogy felkészítsen minket. Együtt próbáltunk neurális hálózatokat szimulálni. De a kísérletek fáradságosnak bizonyultak: a számítógépek lassúak voltak, a programírás pedig sok fejtörést okozott.

Negyedéves hallgatóként, amikor megszállottja lettem ennek a kutatásnak, volt egy megérzésem egy matematikailag nem igazán megalapozott tanulási szabályról a többrétegű neurális hálózatok tanítására. Elképzeltem egy algoritmust, amely a kimenettől visszafelé haladva jeleket továbbít a hálózaton keresztül, hogy a hálózatot így a kimenettől a bemenetig tanítsa. Ezt az algoritmust HLM-nek *(hierarchical learning machine)* neveztem el, és elég büszke vagyok a szójátékomra…[21] A HLM a „gradiens-visszaterjesztéses algoritmus" *(gradient backpropagation algorithm)* előfutára, amely utóbbit ma általánosan használják a mélytanuló rendszerek tanításánál. A HLM nem a gradienseket terjeszti visszafelé a hálózaton keresztül, mint manapság, hanem a kívánt állapotokat juttatja el visszafelé az egyes neuronokhoz. Ez lehetővé tette a bináris neuronok használatát, ami a számítógépek akkori lassúsága miatt a szorzások elvégzésénél előnyös volt. A HLM volt az első lépés a többrétegű hálózatok képzése felé.

A TANULÁS KONNEKCIONISTA MODELLJEI

1983 nyarán megkaptam a mérnöki diplomámat. Rábukkantam egy másik könyvre, amely az önszerveződő rendszerek és automatahálózatok iránt érdeklődő francia kutatócsoport, az LDR (Laboratoire de dynamique des réseaux, azaz Hálózatdinamikai Laboratórium) munkájáról számolt be. Ez egy független kutatóműhely volt, amelynek a párizsi École Polytechnique egykori helyiségei adtak otthont a Rue de la Montagne Sainte-Geneviève-en. A csoport tagjai mind akadémikusok voltak, akik máshol is dolgoztak a felsőoktatásban. Kevés erőforrással, nulla költségvetéssel és egy használt számítógéppel rendelkeztek. Ez is csak azt mutatta, hogy a gépi tanulás kutatása megállt Franciaországban! Meglátogattam őket. Ezek a tudósok nem fedezték fel maguknak a neurális hálózatokról szóló régebbi publikációkat, mint én, viszont ismertek másokat.

Elmagyaráztam nekik, hogy a téma érdekel, és hogy a mérnöki iskolámban van felszerelésem, és csatlakoztam a csoportjukhoz. Közben a Pierre-et-Marie-Curie Egyetemen folytattam a posztgraduális tanulmányaimat. 1984-ben le kellett adnom a doktori disszertációm témaválasztását. Az ESIEE-től társult kutatói ösztöndíjat kaptam, de keresnem kellett egy témavezetőt. Sokat dolgoztam Françoise Fogelman-Souliéval (ma Soulié-Fogelman), aki akkoriban a Párizsi V. Egyetem informatikatanára volt. Logikusan neki kellett volna felügyelnie a doktori disszertációmat, de nem volt erre jogosult, mivel még nem szerezte meg a habilitációját (a habilitáció az európai oktatási rendszer sajátossága).

Így hát megkerestem a labor egyetlen olyan tagját, aki képes volt felügyelni egy informatikai doktori disszertációt: Maurice Milgramot, a Compiègne-i Műszaki Egyetem informatika és mérnöki tudományok professzorát. Ő beleegyezett a dologba, de elmagyarázta, hogy nem sokat fog tudni segíteni, mivel semmit sem tud a neurális hálózatokról. Örökké hálás leszek neki az irántam tanúsított kedvességéért. Az időmet az ESIEE (és a nagy teljesítményű számítógépei), valamint az LDR (és annak szellemi környezete) között osztottam meg.

Ismeretlen vizeken jártam. Izgalmas kaland volt.

Külföldön az enyémhez hasonló kutatások indultak el. 1984 nyarán Françoise Fogelmant elkísértem Kaliforniába, ahol egy hónapos szakmai gyakorlatot töltöttem a Xerox legendás PARC laboratóriumában. Akkoriban az járt a fejemben, hogy két ember van a világon, akivel nagyon szeretnék találkozni: az egyik Terry Sejnowski, a baltimore-i Johns Hopkins Egyetem biofizikusa és neurobiológusa, a másik Geoffrey Hinton, a pittsburghi Carnegie Mellon Egyetem kutatója – ő az, aki velem és Yoshua Bengióval együtt elnyerte a 2018-as Turing-díjat. Hinton és Sejnowski 1983-ban publikált egy tanulmányt a Boltzmann-gépekről, amely „rejtett egységekkel”, azaz a bemenet és a kimenet közötti rétegekben elhelyezkedő neuronokkal dolgozó hálózatos tanulási eljárást vázol fel. Ez a cikk pontosan amiatt nyűgözött le, mivel többrétegű neurális hálózatok tanításáról szólt. Ez áll az én munkám középpontjában is! Ők az én embereim!