

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.*

© ODILE JACOB, 2019

Hungarian translation © Moldvay Tamás, 2025

Hungarian edition © Typotex, Budapest, 2025

Engedély nélkül semmilyen formában nem másolható!

Lektorálta Kertész Gábor

ISBN 978 963 493 335 9

TARTALOM

BEVEZETÉS	7
1 AZ MI-FORRADALOM	11
2 RÖVID ÖSSZEFOGLALÓ AZ MI TÖRTÉNETÉRŐL... ÉS A SAJÁT PÁLYAFUTÁSOMRÓL	21
3 EGYSZERŰ TANULÓGÉPEK	64
4 TANULÁS MINIMALIZÁLÁSSAL, TANULÁSELMÉLET	96
5 MÉLY NEURÁLIS HÁLÓZATOK ÉS A VISSZATERJESZTÉS	134
6 KONVOLÚCIÓS HÁLÓZATOK, A MESTERSÉGES INTELLIGENCIA ALAPPILLÉREI	159
7 A GÉP BELSEJÉBEN, AVAGY A <i>DEEP LEARNING</i> NAPJAINKBAN	186
8 FACEBOOKOS ÉVEK	219

9	MIT TARTOGAT A JÖVŐ? AZ MI ELŐTT ÁLLÓ KIHÍVÁSOK ÉS PERSPEKTÍVÁK	240
10	KIHÍVÁSOK	274
	ÖSSZEGZÉS	302
	KÖSZÖNETNYILVÁNÍTÁS	305
	JEGYZETEK	307

BEVEZETÉS

„*HAL, nyisd ki a légszilip ajtaját!*” A 2001: *Űrodüsszeia* című filmben az űrhajó küldetését irányító szuperintelligens számítógép, a HAL9000 nem hajlandó kinyitni a légszilipet 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 teszük 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 mögöttes meghúzó tudományos eljárást, amely a számítástechnika és az idegtudomány kereszteződésében helyezkedik el. Próbálom ráadásul mindezt úgy tenni, hogy ne pusztá 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 autózvezetést 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 intuíción 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 működesre 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 alapkutatókat. 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 Hintonnal¹ 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 kiszűrják a daganatokat. Az emberi agy azonban jelentős előnyt tudhat a magáénak: sokkalta univerzálisabb és sokkalta képlekenyebb.

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 szerverekre, azok átírják szöveggé, értelmezik, lehívják az információt a meteorológiai 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 Pennsylvanai Egyetemen, hogy a lövedékek röppályáját számítsák ki vele, másodpercenként kb. 360 szorzást végzett tízjegyű számokkal. Ma már ez egy ősközületnek 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 nevekkal.

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 esszéírás-, 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”. Tet- szőleges 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 projektje⁵ a 2019. március 21-én bemutatott, Johann Sebastian Bach születésnapja előtt tisztelgő Google Doodle⁶ 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” Szimfóniájának 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 To-dai. Csak vonzóbb nála a plasztikus megjelenésével, és mi, emberek, akikre ez az animált tárgy hatással van, azt hisszük, hogy valamiféle intelligenciával rendelkezik.