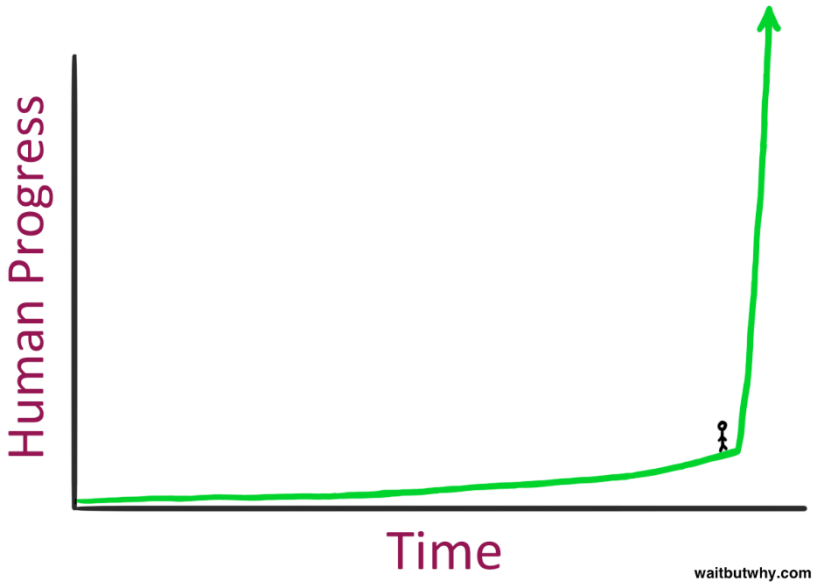
Mesterséges intelligencia és fenntarthatóság: Oxigénhiány a csúcs felé vezető úton

Kapornaki Mihály I Trendency Online Zrt. I 2025.10.13.

A médiát gyakorta a mesterséges általános intelligencia ígérete dominálja: az arról való diskurzus, hogy mi mindenre lesz, vagy legalábbis lehet képes egy-egy csúcsmodell a jövőben. Csak spekulálni lehet, hogy egyszer valóban megalapozottnak bizonyulnak-e majd ezek a remények, de azt már most látjuk, hogy a fejlesztés költségei óriásiak, amelyek fényében a trend töretlensége is megkérdőjelezhető.

A jövő kapujában

Az elmúlt évek technológiai diskurzusát egy képbe sűrítve – a mesterséges intelligencia által hajtott exponenciális fejlődés ígérete – festhetjük le.



A forradalmi technológia nem szükségszerű, hogy forradalmi alkalmazásokhoz vezessen. Az „AI" címke mára sokat inflálódott, azok miatt, akik egy-egy színjátékhoz, vagy elhamarkodottan használták. A kor szlogenje a „Powered by AI" lehetne, de hogyan nézett ki ez az innovációs forradalom a mindennapokban?

Az Amazon Fresh üzletei „Just Walk Out" vásárlási élményt lehetővé tevő technológiát hirdettek: nézelődünk, kiválasztjuk a kedvünkre való termékeket, az intelligens kamera felismeri, hogy mit raktunk a kosarunkba, a levonás pedig azután érkezik, hogy kisétáltunk a boltból – nincs többé sorban állás a kasszáknál. Gyors és kényelmes, még ha a számla egy automata rendszerhez képest olykor meglepően sokára jött is meg. Sajnos az élmény valódi motorja nem az MI, hanem nagyjából ezer, főként Indiából származó munkavállaló volt, akik aktív közreműködése kellett a tranzakciók több mint 70%-ához.

Elszigetelt eset, mondhatnánk, – még ha furcsán is hangzik a világ egyik legnagyobb vállalata kapcsán – de a közvéleményt időről-időre felkorbácsolják az olyan esetek, amikor egy technológiát többnek adnak el, mint ami, ahogyan az a Builder.ai esetében is történt.

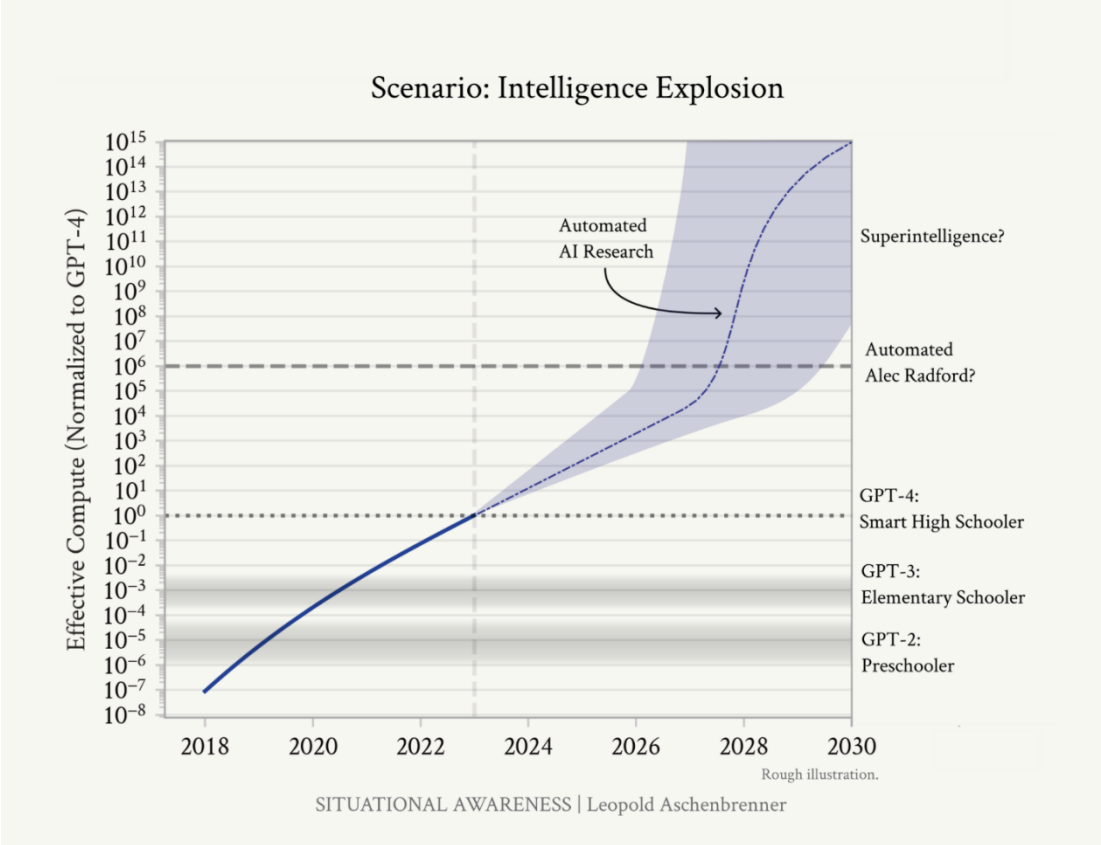
A csúcson másfél milliárd dollárra értékelt londoni startup azt ígérte, hogy Natasha nevű chatbotjuk egy app vagy website 80%-át elkészíti pusztán szöveges leírás alapján (fejlesztői tudás nem szükséges!). A vállalat hitelét azonban meglehetősen lerontotta, – azon túl, hogy az „AI-assisted" és az „AI-generated" közötti különbségtételre egyre többen kezdtek el figyelni – amikor kiderült, hogy a cég számlázási csalással érte el, hogy nyereségesebbnek tűnjön, mint ami.

De ha mi magunk nem is vettünk igénybe ilyen szolgáltatásokat, van olyan, nap mint nap elénk kerülő funkció, mint az internetes keresők összefoglalói (AI Overview), amelyek rávilágítanak egyebek mellett a megfelelő tanítóadat (az emberek esetében a forráskritika) fontosságára, például amikor felhívják a figyelmünket a sziklák helyére az egészséges étrendben, vagy azt tanácsolják, hogy a túl könnyen lecsúszó pizzafeltétre a ragasztó hozzáadása a megoldás.

Milyen jelentőséget tulajdonítsunk ezeknek az eseteknek, milyen fejlődési ívre illeszthetők rá?

Leopold Aschenbrenner (2024) nagyívű esszésorozata az évtized végére már lehetségesnek tartja az emberi értelmet nagyságrendekkel meghaladó szuperintelligencia létrejöttét. Az út első szakasza a számítási kapacitás és algoritmikus hatékonyság növekedése és a modellek látens képességeinek kibontakoztatása, például azáltal, hogy eszközöket használva chatbotból ágensekké válnak. Ez vezetne az általános mesterséges intelligencia (AGI) létrejöttéhez (szerinte akár már 2027-re!), amely már a legjobb emberi szakértők szintjén van, és kellően nagy számban képes az MI kutatást, mint olyat automatizálni, amely rekurzív

önfejlesztési hurokként működve olyan robbanásszerű képességjavulást eredményez, amely egy évtizednyi fejlődést kevesebb, mint egy évbe sűrít.



Vegyük azonban észre, hogy az ilyen és ehhez hasonló víziók mindig inkább a potenciálról szólnak. Tehát itt egy jelentős szakadékot látunk a Phd-szintű (sőt szuper-) intelligencia teremtette lehetőségek és a jelenlegi hasznosulás között. Az átkeléshez pedig kolosszális befektetések kellenek, miközben a siker egyáltalán nem magától értetődő. A továbbiakban a legnagyobb akadályokat emeljük ki empirikus példákon keresztül: az adatot és az energiát.

MI-beltenyészet

Intő jelként szolgálhat az, ahogyan Guo és munkatársai (2023) a nyelvi diverzitás csökkenését mutatták ki a generatív technológia kapcsán. A módszerük alapja, hogy a tanítást rekurzív módon szimulálták, vagyis az első bázismodellt (OPT-350M) még autentikus, feladatspecifikus adatokkal finomhangolták, míg a következő modellek az előző modell által generált szintetikus adatokon tanultak – 6 iteráción keresztül.

A modellek különböző nyelvi feladatokat végeztek angolul: hírek összefoglalása, tudományos absztrakt létrehozása, történetírás (behatárolttól a kreativitást igénylő tevékenységig).

Eredményeik a diverzitási metrikák konzisztens csökkenését mutatták az iterációk során, tehát kevesebb, kevésbé változatos kifejezés (pl. szinonimák), egyszerűbb mondatszerkezetek (pl. alá-fölérendeltségi viszonyok komplexitásának csökkenése) és kevesebb témakör. A romlás a high-entropy (kreatívabb) feladatokban volt a legmeredekebb és különösen látványos volt a szintaktikus sokszínűség vesztesége.

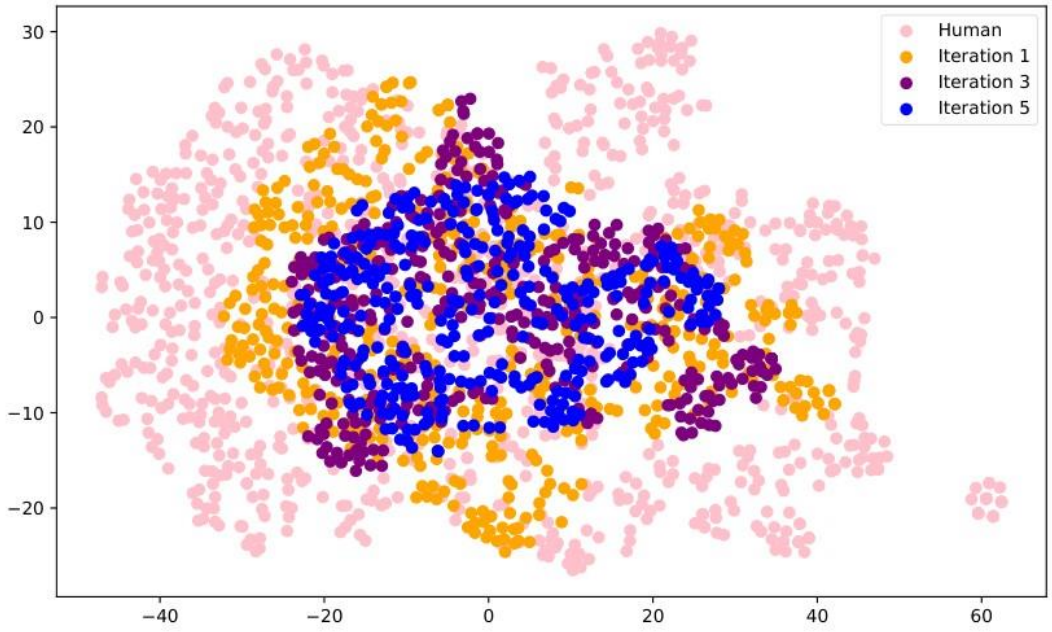


Figure 4: T-SNE visualization of dependency tree embeddings derived from sentences generated in successive iterations of our tuning-generation process. The visualization clearly depicts how, over time, the spatial distribution of the embeddings becomes increasingly compact. This decreasing spread is indicative of declining syntactic diversity.

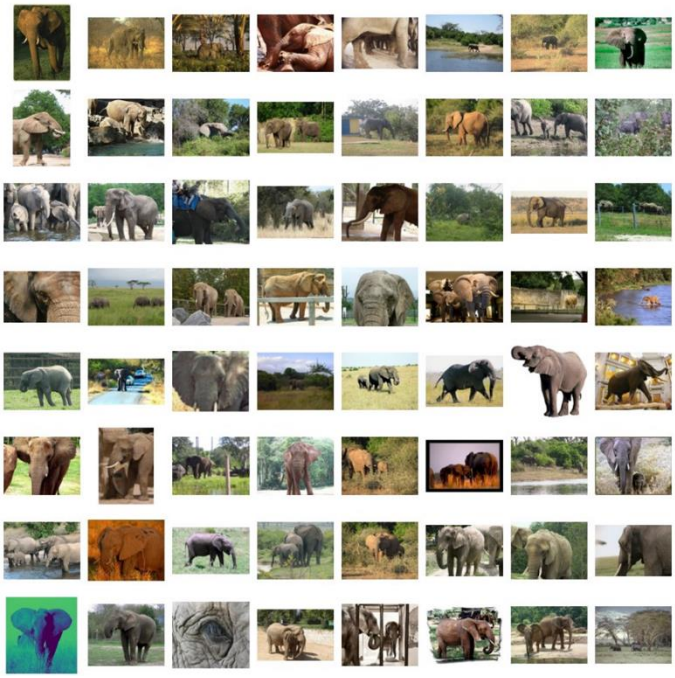
A fenti ábra a szintaktikus veszteséget szemlélteti: láthatjuk az emberi mondatok szóródását a teljes térben, szemben a fokozatos iterációk (szintetikus adatok) következtében létrejövő kompaktsággal.

Jusson eszünkbe, hogy a kereskedelmi modellek elterjedésével eleve növekvő mennyiségű modell által generált adat jön létre és kerül a jövőbeli típusok elé. Ahogy a jó minőségű emberi adat fogyatkozik, úgy a fejlesztőknek nem is nagyon lesz más választásuk, minthogy ezekre a „mesterséges” adatokra támaszkodjanak. A szerzők kísérlete így nagyon is valós veszélyre mutat rá.

Természetesen ez a problémakör nem csak a szöveges adatokat érinti. Hataya, Bao és Arai (2023) szó szerint látványos példát tárnak elénk a képi adathalmazok kontaminációját szimulálva. StableDiffusion segítségével nagyméretű adathalmazokat alkottak ImageNet és COCO kategóriák számára. A generált képeket valódi adatok közé keverték változó arányban (20%, 40%, 80%) és különböző feladatokon (klasszifikáció, feliratozás, képgenerálás) teszteltek. Eredményeik kiterjedt teljesítményromlást mutattak, legfőképpen nagy kontaminációs szinten, kiemelten a képek osztályozása terén, amelyben akár 60 százalékpontos romlás is kimutatható volt.

Az okot beláthatjuk, hogy ha lentebbi kollázsokat összehasonlítjuk: az eredetihez képest, a generált képek lényegesen kisebb változatossággal bírnak (az adatpontok kevés nagy sűrűségű klaszterbe esnek).

Eredeti



Generált

A szintetikus adatok felhasználhatósága nagyon is komoly gyakorlati kérdés, egyrészt, mert a mikroműanyagokhoz hasonlóan mindenhol jelen vannak, másrészt mert a mesterséges intelligencia valójában kollektív intelligencia (a felhalmozott digitalizált emberi tudásból tanult), de az internet bugyrai sem végtelenek (főleg, ha minőségi adatot akarunk) és már látszik, hogy nem születik annyi új tartalom, amilyen ütemben azt a modellfejlesztés igényelné. Márpedig, ha nem tartjuk kontroll alatt, akkor a szintetikus adatokon való tanulás a szivacsos agyhártyagyulladáshoz lesz

hasonlatos – amolyan digitális kergemarhakórhoz, amikor az állat (esetünkben a modell) saját fertőzött fajtársa által szennyezett eleségtől betegszik meg. A probléma tehát egyáltalán nem marginális, mert újabb és még nagyobb modellek tanításához kolosszális méretű adathalmazok kellenek.

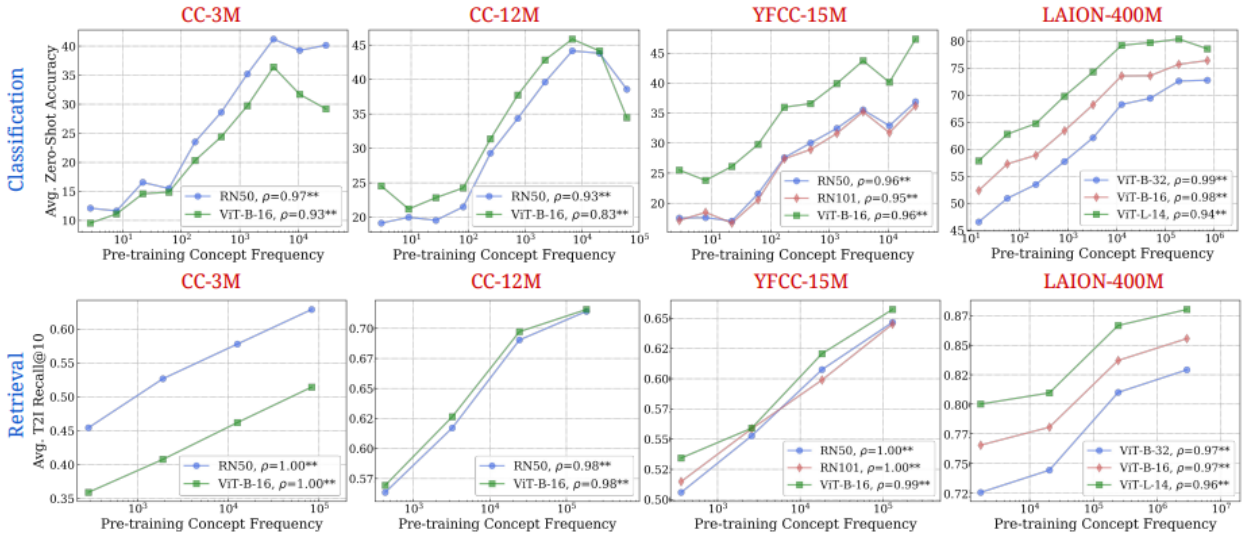
Adatigény

Ugyanakkor a több adat = jobb modell összefüggésnek is megvannak a korlátai, ahogyan arra Udandarao és munkatársainak (2024) rendkívül alapos és szemléletes vizsgálata rá is mutatott. A kutatók 34 multimodális modellt, mint amilyen a képi és szöveges adatokon tanított CLIP, vagy éppen a StableDiffusion, teszteltek 5 standard nagy méretű adathalmazon.

A kutatási kérdés arra irányult, hogy hogyan befolyásolja a különböző fogalmak, konkrét objektumok vagy kategóriák pl. férfi, aranyhal kalap stb., gyakorisága a tanítóhalmazban a modellek teljesítményét. Ennek érdekében három főbb feladattípust mértek:

- osztályozás: adott kép besorolása egy kategóriába,
- visszakeresés: egy képhez felirat társítása, illetve fordítva, valamint
- képgenerálás: minőség és illeszkedés mérése

Mindhárom esetben a teljesítmény a fogalmaknak a tanítóadatban való gyakoriságának függvénye volt.



A szerzők konzisztensen azt találták, hogy a multimodális modellek messze járnak a zero-shot általánosítástól, ellenben exponenciálisan több adatra volt szükségük, hogy lineáris javulást érjenek el a feladatokmegoldását illetően, a minta nem hatékony log-lineáris skálázási trendjét követve.

Itt érdemes megállnunk egy pillanatra. A tanulmányban is használt „zero-shot" tanulás kifejezés lényegében a valódi általánosítást jelentené – a megtanult mintázatok megértésének és új esetekre való alkalmazásának képességét.

Hagyományos értelemben tehát a modell jól teljesítene a tanítás során még nem látott fogalmak esetében is: felismerne egy macskát, hiába nem látott róla képet vagy képes lenne zebráról képet generálni kompozíciós megértés alapján (csíkok + lóhoz hasonlatos állat).

Valódi zero-shot esetén tehát a teljesítménynek függetlennek kellene lennie a tanítási gyakoriságtól, nem pedig közvetlenül korrelálnia vele, ahogy a tanulmány mutatja. Itt tehát a modellek intelligens voltát túlhangsúlyozó narratívával ellentétben nem kifinomult következtetést, hanem rendkívül erőteljes, de hatékonytalan memorizációt látunk!

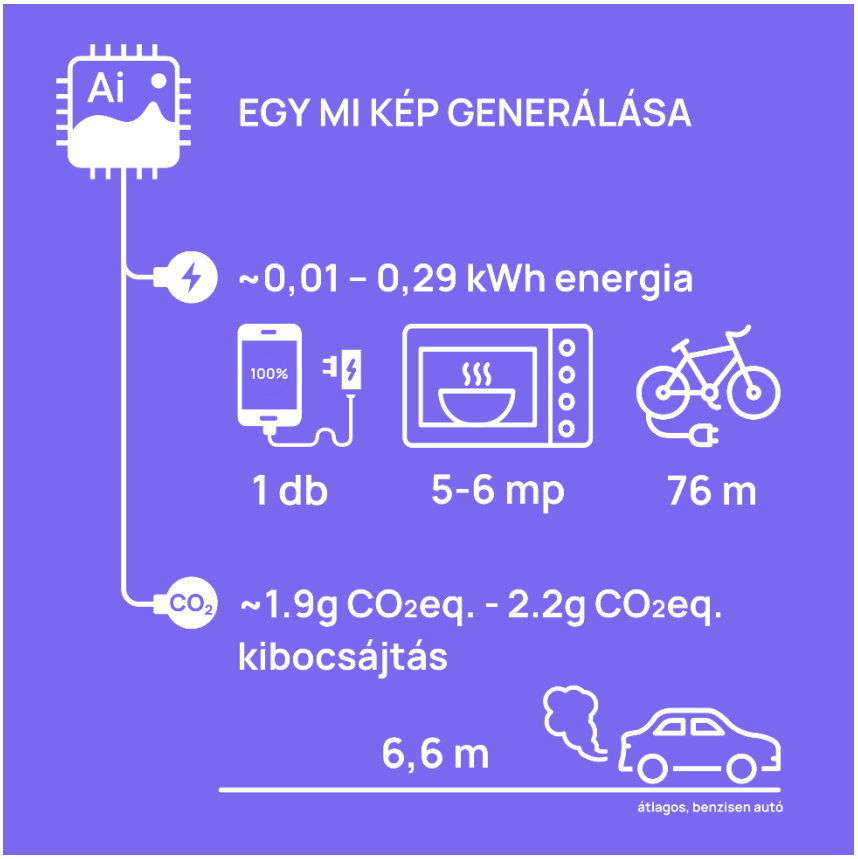
Energiaigény

A kilátások borúsabbak lesznek, ha belegondolunk, hogy az eddigi modellek sem csak úgy felszívták az adatot, majd működtek maguktól, hanem mint a kifinomult gépeknek általában, áramra volt szükségük. Ezt hangsúlyozta Luccioni, Jernite és Strubell (2024), amikor 88 feladatspecifikus és többcélú modellt vizsgáltak energiaigény és karbonkibocsájtás szempontjából benchmark adathalmazokon tíz különböző feladaton keresztül (pl. képklasszifikáció, szövegalkotás, objektumészlelés).

A tanulmány több fontos megállapítást is tartalmaz. Láthatóvá tesz egyfajta modalitási büntetést, vagyis, hogy egy feladat típusa többet számít, mint egy modell mérete: egy képpel kapcsolatos feladat 3-60x annyi energiát igényel, mint egy szöveges feladat, akkor is, ha kontroll alatt tartjuk a modellparaméterek számát. Megfigyelhető továbbá a generatív-diszkriminatív trade-off: értjük ezalatt, hogy a tartalmat (szöveg, kép) előállító feladatok nagyságrendekkel drágábbak, mint a létező tartalom klasszifikációja, kategorizációja. Ami a mondanivalónk szempontjából különösen hangsúlyos, az egyfajta reality-check a nagyságrendeket illetően.

Számos beszámoló csak a modellek betanításának energiaköltségét (és/vagy karbonlábnyomát) hangsúlyozza, ami valóban látványos, ahhoz képest, amikor a modellt már csak működtetni kell (inferencia). Csakhogy a tanulmány világosan kimutatja, hogy a népszerű modellek esetében a költségparitás az interferenciák nagy száma miatt gyorsan elérhető. Arról van itt szó, hogy a nagy kereskedelmi modelleknek a csúcsnapjaikon annyi felhasználójuk van, hogy csupán 1 query/felhasználó esetén is, heteken, de legfeljebb hónapokon belül meghaladja a használat a betanítás energiaköltségeit.

Ez pedig rendkívüli terheket ró ránk.



Szigorúan szakmai szempontból fontos üzenet lehet itt, hogy az általános célú modellek nagyon erőteljesek, – amennyiben sok feladatot és jól hajtanak végre, – de nagyon nem hatékonyak vagy gazdaságosak – a hagyományos gépi tanulási megoldásokhoz képest relatíve lassúak és drágák. Érdemes megfontolni, hogy ahelyett, hogy mindenre ezeket eresztenénk rá, sokszor racionálisabb, ahol lehet feladatspecifikus modelleket használni. Az energetikai és környezetvédelmi szempontok nem a mi szakterületünk, de ha belegondolunk, hogy naponta

megközelítőleg 34 millió képet generálunk, beláthatjuk, hogy a hatékonyság kérdésének tétjei óriásiak.

Márpedig ebben a kérdéskörben kiemelten fontos, hogy a jövőnk szempontjából is gondolkodjunk. A Nemzetközi Energiaügynökség (IEA) 12%-os növekedést mutatott ki az elmúlt 5 évben az adatközpontok áramfogyasztásában. Előrejelzéseik szerint 2030-ig évente15%-os növekedés várható, ami négyszerese a többi szektorénak együttesen.

Ez a trend kiemelt figyelmet érdemel, lévén az adatközpontok már így is a globális áramfogyasztás hozzávetőlegesen 1,5%-át adják (415 TWh/év), amely mögött az MI-fejlesztés a legfőbb hajtóerő. Ez nagyságrendileg Franciaország éves áramfogyasztásának felel meg. Ha az IEA előrejelzései pontosnak bizonyulnak, akkor ez a szám 2030-ra 945 TWh lesz évente, amely már a globális fogyasztás 3%-át (Japán jelenlegi fogyasztását) tenné ki.

A korábban idézet Aschenbrenner jóslatai alapján (l. alábbi táblázat) az általa szuperintelligenciának titulált modellnek csak a betanítása 100 nagy nukleáris reaktor kapacitását igényelné és több mint 1 billió (vagyis ezermilliárd) dollárba kerülne. A nagyságrendekkel erősebb modellek tehát a rengeteg pénz mellett forradalmi áttörést igényelnének a chipgyártásban és energetikában is.

Year	OOMs	# of H100s-equivalent	Cost	Power	Power reference class
2022	~GPT-4 cluster	~10k	~$500M	~10 MW	~10,000 average homes
~2024	+1 OOM	~100k	$billions	~100MW	~100,000 homes
~2026	+2 OOMs	~1M	$10s of billions	~1 GW	The Hoover Dam, or a large nuclear reactor
~2028	+3 OOMs	~10M	$100s of billions	~10 GW	A small/medium US state
~2030	+4 OOMs	~100M	$1T+	~100GW	>20% of US electricity production

Természetesen az MI és a tágabban vett gépi tanulás nem csak a problémának, hanem a megoldásnak is a részese lehetne. A DeepMind prediktív modellje például lehetővé tette a Google adatközpontjainak hűtőrendszereinél az energiaköltségek 40%-os csökkenését. A probléma, hogy ez a bíztató példa 2016-ból származik, azóta pedig rengeteg karbon-semlegességgel, zöld-energiával, hatékonysággal foglalkozó kutatást visszafogtak, célkitűzést eltöröltek, mert kellett a forrás a generatív modellekre és az őket kiszolgáló infrastruktúrára.

Mihez kezdjünk mindezzel?

Az MI tündöklése terebélyes árnyékot vet, de meg kell tanulnunk alkalmazkodni a világossághoz és a sötétséghez egyaránt. Ezt az írást az „imádkozz, de tartsd szárazon a puskaport" hozzáállása vezette.

Ha megoldjuk a korábban jelzett problémákat - energetikai forradalommal (pl. fúziós energiával), gazdaságosabb számításokkal, nagyobb emberi közreműködéssel és/vagy kifinomultabb szintetikusadat-generálással, - ha nem, nekünk többieknek, akik nem közvetlenül nagy alapmodelleket fejlesztünk, készen kell állnunk a mesterséges szuperintelligencia előtt és nélkül is jól működni és értéket teremteni. Nézzünk ehhez néhány kapaszkodót!

Open source

Amikor megfelelően biztonságos, jól teljesítő megoldás áll rendelkezésünkre, akkor igyekezzünk nyílt forráskódú programokra támaszkodni, hogy csökkentsük a kitettségünket, ha a nagy piacvezető projektek megakadnának – emellett járuljunk mi is hozzá ennek az eszköztárnak a gyarapodásához.

Szabályozás

Erre egy példa az EU AI Act. Nem teszünk kísérletet megjósolni, hogy egy-egy szabályozás milyen hatással jár (működik-e még a híres Brüsszel-hatás), de az biztos, hogy a technológiai változások milyenségét továbbra is az fogja meghatározni, hogy társadalomként hogyan reagálunk rá. Ha azt várjuk az MI-től, hogy fenekestül felforgatja a világot, akkor azt jogilag is felkészülten kell várnunk.

Szilárd szakmai és szervezeti alapok

Bármilyen technológiát is építünk be a termékeinkbe és/vagy a munkafolyamatinkba, annak hozzáadott értékét csak úgy tudjuk igazán kihasználni, ha az alapokat már

lefektettük a tiszta kód, a hatékony projektmenedzsment, a jó ügyfélkommunikáció, a mindezt támogató szervezeti kultúra és a mély szakértelem révén. Ezek megléte jellemzi azokat a cégeket, akiknek valódi keresnivalója van kísérleti technológiák integrálásában.

Kétkedés

Végül pedig kiemelten fontossá válik a megfelelő kérdések folyamatos feltétele: bármit is harangozzanak be a hírek, vizsgáljuk meg, hogy mik azok az elvek és eszközök, amik MI utópia és disztópia esetén is egyaránt segíteni fognak minket és hogy mi az, amit már MA jól tudunk csinálni a rendelkezésünkre álló technológia és szervezeti eszköztár segítségével.