

Az emberi intelligencia kiterjesztése - hogyan használjuk a mesterséges intelligenciát?

Arzberger László

arzbergerlaszlo@proton.me

Öntudatra ébredhet-e a mesterséges intelligencia? Vagy csak egy újabb technikai lufi az egész MI-láz? Eldönti-e az MI, hogy milyen gyógyszer szedjek? Mik az etikai korlátok? Cikkünkben ennek a technológiának a korlátait és megbízható használatának a lehetőségeit vizsgáljuk meg a szenzációkeresés mellőzésével, a tudomány mai eredményeire támaszkodva.

Az OpenAI nagy nyelvi modellje, a ChatGPT 2022 novemberi nyilvános bemutatásával elkezdődött technológia forradalom alapvetően változtatja meg életünk több szegmensét, olyan megoldásokkal és folyamatokkal, amelyekkel eddig legfeljebb a sci-fi kedvelői találkozhattak a könyvek lapjain vagy a filmvászonon. Szinte naponta jelennek meg szakmai és kevésbé szakmai jóslatok a mesterséges intelligencia jövőjét és felhasználását illetően. Itt pontosabban a generatív mesterséges intelligenciára gondolok, amely jellemzője, hogy szöveges, képi, hangi, vagy videós tartalmat hoznak létre megadott instrukciók alapján. Jellemzően kódra sincsen szükség, ezekkel a modellekkel emberi nyelven lehetséges kommunikálni: különösen igaz ez az LLM-ek, a nagy nyelvi modellek esetében. Egyesek már az öntudatra ébredt technológia rémképét vizionálják, mások szerint az MI rövidesen olyan alapvető módon lesz mindennapjaink része, mint az áram vagy az internet. Megint mások a generatív mesterséges intelligencia megbízhatatlanságát emlegetik: laikus felhasználók gyakran panaszkodnak, hogy az MI „hazudik”. Egyesek szerint az egész MI-láz csak egy technikai lufi, amit jó nagyra fújtak.

A legfontosabb rögtön leszögezni, hogy az emberi nyelven történő kommunikáción és emberi gesztusok imitálásán túl ezek a modellek a betanított információk által generálnak új tartalmakat. Az utasítások, a „promptok” pedig bár emberi nyelven vannak, a pontos és korrekt változathoz elengedhetetlen a különböző promptképző metódusok ismerete. A felhasználás korlátja sok esetben éppen ez. Mivel betanított adatok alapján dolgoznak, nem képesek erkölcsi és etikai alapú döntéshozatalra, és az adatokban rejlő előítéletek alapvetően befolyásolni tudják a generált választ. Éppen ezért nem alkalmasak például egy pályázat elbírálására, bírói döntésekre vagy arra, hogy milyen intézkedés szükséges vészhelyzetben, például indokolt-e vagy sem mentőt küldeni valakihez. Ugyancsak aggályos és veszélyes a kritikus infrastruktúra (víz- és energiaellátás, közlekedés), valamint automatizált fegyverrendszerekben való alkalmazása, ha a végső döntést nem ember hozza meg. Itt fontos megemlíteni az adatvédelem, és az azzal kapcsolatos törvényi szabályozások tiszteletben tartását is.

Nagy hangsúlyt kaptak a híradásokban az MI öntudatával kapcsolatos viták és kutatások. 2023 augusztusában egy 19 informatikus kutatóból, idegtudósokból és filozófusokból álló csoport [létrehozott egy hosszú ellenőrzőlistát](#), amelyek utalhatnak a mesterséges intelligencia tudatosságára, de nem feltétlenül bizonyítják azt. Eredményeiket egy hosszú vitairatban publikálták, amiben az emberi tudat-teóriák alapján 14 kritériumot ajánlanak, majd azokat alkalmazták az elérhető MI-architektúrákra, többek között a ChatGPT modellre is. Arra a következtetésre jutottak, hogy valószínűleg egyikük sem tudatos. Robert Long, a San Francisco-i nonprofit Center for AI Safety munkatársa szerint ez keretet ad az egyre emberszerűbb mesterséges intelligencia kiértékeléséhez. „Olyan szisztematikus módszertant vezetünk be, amely korábban hiányzott.” Fontos megjegyezni, hogy az összes ilyen projekt az emberi tudatról szóló ismereteken alapul, azonban a tudat egészen más formákat ölthet, akár emlőstársainkban is. Ahogy az egyik résztvevő Adeel Razi fogalmazott: „Igazából fogalmunk sincs, milyen denevérnek lenni. Ez egy olyan korlát, amelytől nem tudunk megszabadulni.”

A generatív mesterséges intelligencia használatának kockázata, hogy helytelen válaszokat adhat, ami jelentős károkhoz vezethet, különösen olyan vitális területeken, mint az egészség és a biztonság. Erre a nem megfelelő felhasználásra jó példa lehet, ha megkérdezzük, hogy biztonságos-e egy adott gyógyszer alkalmazása anélkül, hogy az adott illető egészségi állapotáról is információt adnánk.

A generatív MI megfelelő alkalmazása az, amikor közvetlen válaszok keresése helyett a felhasználó a kérdéses terület jobb megismerésére törekszik. Például megkérdezi az MI-t, hogy ha orvoshoz készül egy adott problémával, ott milyen kérdéseket tegyen fel ezzel kapcsolatban. Ez a megközelítés minimálisra csökkenti a kockázatot azáltal, hogy biztosítja, hogy egy képzett szakember hozza meg a végső döntést, aki hozzáfér minden szükséges információhoz.

Kulcsfontosságú, hogy a generatív mesterséges intelligencia használata olyan módon történjen, amely emberi felügyeletet igényel, ezzel csökkentve a károk lehetőségét, de ugyanakkor előnyöket is biztosít. A párbeszéd és kérdések felvetése jobb megértést és döntéshozatalt eredményezhet, biztosítva, hogy az MI-eszközök használata felelősségteljes és a felhasználó számára előnyös legyen.

A közelmúltban zajlott egy érdekes kísérlet az [utrechtii egyetem szervezésében](#), ahol MI-kutatók a világ minden részéről érkezett művészekkel kutatták annak a művészetekre gyakorolt hatását. Képzőművészek, fotósok, zenészek alkotásait módosították generatív MI-vel, aztán ezt a módosítást kellett reprodukálni a résztvevőknek. A projekt megmutatta, hogy ezek a rendszerek értékes eszközökként szolgálhatnak az alkotói folyamatban, de nem helyettesíthetik a művészeket és az emberi kreativitást.

Itt kell rögtön rátérnünk a technológia egyik sajátos jellemzőjére, a hallucinációra, amire nem hibaként, hanem egy olyan sajátosságként érdemes tekinteni, ami jelentősen kiterjeszti a felhasználási lehetőségeket, például ötletek generálásához. Több szakember azt a megközelítést alkalmazza, hogy a mesterséges intelligencia helyett jobb a kiterjesztett intelligenciáról beszélni, mert a cél nem az emberek kiváltása, hanem az

emberi elme “kiterjesztése” olyan eszközökkel, amelyek javítják az emberi kreativitást, problémamegoldó-képességet, és termelékenységet. Használhatjuk például az MI-t arra, hogy segítse a csoportokat a jobb koordinációban jegyzőkönyvek összegzésével, a tervek kétértelműségének tisztázásával, és a lehetséges konfliktusok azonosításával. Automatizálhatjuk az ismétlődő és unalmas hétköznapi munkát MI-vel, hogy az emberek értelmesebb és vonzóbb tevékenységekre összpontosíthassák idejüket. Létrehozhatunk egyfajta „biztonsági hálót”, ahol mesterséges intelligencia felülvizsgálja az emberi produktumokat, hogy felismerje azokat a hibákat, amelyeket az ember nem vesz észre. Segíthetjük a problémamegoldást, és az MI-válaszait felhasználva ösztönözhetjük az új ötleteket és a kritikus gondolkodást. Fontos, hogy az MI válaszaira kiindulópontként, és ne végső megoldásként tekintsünk.

Ahogy tehát a fentiekből világosan láthatjuk, az MI használata rengeteg új lehetőséget és előnyt hozhat, ha megfelelően és felelősségteljesen használjuk. A végső döntés viszont mindig a miénk, embereké kell legyen!