A MESTERSÉGES INTELLIGENCIÁN ALAPULÓ ALKALMAZÁSOK

hosszú távon pozitív hatást fognak kifejteni a társadalomban – interjú Szabó W. Péter fejlesztővel

W. Szabó Péter informatikus a digitális platformok felhasználói élményének kutatásával kezdte karrierjét, majd a mesterséges intelligencia felé fordult a figyelme. Tavaly novemberben indította el hivatalosan a Tengr.ai *nevű képgeneráló szoftvert, amely rövid időn belül rendkívüli népszerűségre tett szert.*
Szenvedélyesen hisz abban, hogy a negyedik ipari forradalom, amely jelenleg világszerte zajlik, hosszú távon pozitív hatással lesz a társadalomra.

■ – *A mesterséges intelligencia kifejezés helyett inkább a gépi tanulás fogalmat szereted használni. Mit jelent a gépi tanulás?*

– A mesterséges intelligencia fogalma nagyon általános, ezért sokan használják mindenféle kontextusban anélkül, hogy valójában tudnák, mit jelent. Ráadásul már a marketing is a végtelenségig elkoptatta, főleg amióta a ChatGPT különféle változatai megjelentek. A mesterséges intelligencia annyi mindent jelenthet, annyi területen alkalmazható, hogy pontosítani kell a fogalmat. A gépi tanulást egzakt módon meg lehet határozni: egy konkrét feladatra – például szöveg generálására – betanított modell azért tanul, hogy minél jobb legyen, több hibát elkerüljön. Ugyanúgy ahogyan az ember a gyermekkora óta tanul, a gépi modell is folyamatosan tanul, a visszajelzések alapján kiküszöböli a hibákat, és egyre jobban fog teljesíteni. Persze nullára soha nem lehet csökkenteni a hibalehetőséget.

A távolabbi jövőre nézve mindenképp optimista vagyok. Úgy ahogy a korábbi technikai forradalmak esetében történt, a jelenlegi alkalmazások is rengeteg pozitív dolgot fognak eredményezni, könnyebb lesz a következő generációk élete...

– Hogyan tudjuk megtanítani a gépet arra, hogy elvégezzen egyes feladatokat?

– Jelenleg mindenki az úgynevezett előre tanított modelleket részesíti előnyben. A ChatGPT betűszója is, ha kibontjuk, azt jelenti, hogy generative pretrained transformer, tehát előre tanított generatív átalakító. Ebből az is adódik, hogy az ismeretanyag, az adott szövegtörzs például, amiből tanult egy modell, egy adott dátumig tart, amikor lezárták az adatbázis létrehozását. A GPT–3.5 Turbo esetében azért volt ez a dátum 2021 szeptembere, mert nem sokkal ez előtt jött ki az Nvidia DGX architektúrájába való legújabb videókártya, ami mérföldkőnek számított, és mindenki ennek a segítségével kezdte el tanítani a gépi modelleket. Miután a GP–3.5 modell megtanulta az adathalmazból, hogy miként generáljon szöveget, 2022 novemberében át lehetett adni a közönségnek az előre betanított modellt, amit világszerte kisebb csodának tekintettek, és ezért nagy médianyilvánosságnak örvendett, mindenki forradalmi áttörésről beszélt. Addig a gépi tanulás csak a szakmai konferenciák technikai körében volt beszélgetési téma, és éppen ez az áttörés, egy jó megoldás vezetett oda, hogy elindult a gépi tanulásnak a népszerűsítése, alkalmazása szélesebb körben.

– Hogyan képzeljük el a gépi tanulást? Miből és hogyan tanul a mesterséges intelligencia modellje?

– A tanuláshoz elengedhetetlen, hogy létrehozzunk egy adatbázist, amelyben nagy mennyiségű, megfelelően előkészített adat található. Én sokkal többet foglalkoztam a képgeneráló modellekkel, mint a nyelvi modellekkel, tehát erről tudok részletesebben beszélni. Ahhoz, hogy egy adatbázisból tanuló adathalmaz legyen, szükséges egy adatfelfedezési munka, ami azt jelenti, hogy a modell, tegyük fel a képgeneráló szoftver, végigjárja az adott halmazt, a benne található képeket. A modell egy nagy internetes adatbázisból indul ki, ami képek esetében lehet például a LAION adathalmaz, amiben jelenleg kb. 6 milliárd kép van. Ezek a képek fel vannak címkézve, tehát van egy szöveg, ami leírja, hogy mi látható a képen (például „macska ül az asztalon”).

A szöveggeneráló modellek szintén nagy adathalmazokból tanultak, ez esetben a Pile az adathalmaz, amit jelen pillanatban használnak. A GPT–3-ig biztosan ez volt a kiindulópont, de valószínűleg a GPT–4-es is ezt használja. A későbbi modelleknél már nem egyértelmű, hogy miből tanulnak a modellek, az OpenAI szakmai titoknak minősítette ezen információt. Az eredeti *Pile* adathalmaz kb. 800 gigabyte-nyi nyers szöveg, ami 22 kisebb részből áll össze. Az adathalmazban rengeteg könyv, folyóirat, weboldal szövege, chatbeszélgetés található. A cél az volt, hogy minél változatosabb szövegekből tanulhassa a szöveggenerálást, habár sok esetben merültek fel szerzői jogi kérdések, pl. olyan szerzők művei esetében, akik még élnek, vagy fizetős tanulmányok és szakcikkek használata esetében.

– A szövegek generálása gépi tanulással nagyon sok területen használható. Milyen következményekkel járt ez a médiában?

– A nagy nyelvi modellek úgy tudják támogatni a médiában dolgozókat, hogy az újságíró szolgáltatja a tényeket, és a nyelvi modell azt a feladatot kapja, hogy ennek alapján jó, olvasmányos szöveget írjon. Itt nem lehet az a cél, hogy a gépi tanulás által a nyelvi modell információkat nyújtson, hiszen ezeknél a modelleknél a tanulás azt jelenti, hogy minél kevesebb nyelvtani hibát kövessen el, minél jobb legyen a mondatok szerkezete. A generált szövegek igazságtartalma attól függ, hogy mennyire pontos adatokat szolgáltatunk neki. És éppen amiatt, hogy a nyelvi modell nagyon jó szöveget próbál generálni, megtörténhet, hogy

elkezd „hallucinálni”, ami megszokott jelenség a szöveget generáló mesterséges intelligencia esetében, mivel a szöveg igazságértéke másodlagos. Ha nincs az adathalmazban semmi, amiből az igazságértéket meg tudná saccolni, akkor mond valamit, ami szövegszinten hihető.

Amit aránylag magas szinten meg tud csinálni a nyelvi modell, hogy mondjuk két angol nyelvű híranyagból legyártja a magyar nyelvű hírt. És ezt töredék idő alatt megcsinálja ahhoz képest, hogy a szerkesztő mennyi ideig dolgozott volna rajta. Persze az eredményt minden esetben ellenőrizni kell, mert csúszhatnak be hibák.

A másik negatív tendencia, ami már erőteljesen jelen van, és ami nem volt eredeti célja a nagy nyelvi modellnek, az álhírek generálása. Komoly probléma lett, hogy hihetetlen mennyiségben generálnak álhíreket, ami alkalmas lehet lejáratókampányokra is. Mivel ily módon rengeteg úgynevezett hírforrást elő lehet állítani, az emberek hajlamosak hinni ezeknek az híreknek, annál is inkább, mert a közösségi médiában léteznek a visszhangkamrák, ahol folyamatosan az általunk kedvelt tartalmakkal bombáznak a platformok algoritmusai, és teljesen elzárják az ellentétes tartalmakat.

Az álhírek gyártása úgy történik, hogy valaki valamilyen érdekből kitalál egy alapnarratívát, amit el szeretne ültetni az emberek fejében, majd a nyelvi modell ebből rengeteg hozzá kapcsolódó anyagot gyárt. A koronavírus idején pl. rengeteg összeesküvés-elmélet terjengett a közösségi médiában. Nehéz persze egyértelműen megállapítani, hogy mely szövegeket gyártott a mesterséges intelligencia, de például volt olyan eset, hogy az egyik ismert álhírcsatornán ugrásszerűen jobb lett a helyesírás, hirtelen jobb szövegeket írtak, részletesebb hírek keletkeztek. Ebből gyanítható, hogy elkezdtek mesterséges intelligenciával generálni szövegeket.

Ugyanakkor vannak bothálózatok, amelyek a közösségi médiában is károkat okoznak: feltörnek vagy ellopnak valós oldalakat, és ezeket különféle célokra használják, akár álhírek terjesztésére vagy olyan linkek küldésére, amelyek megfertőznek egy számítógépet. Ezek már komoly cyberbiztonsági kockázatok.

– A képgeneráló programok már majdnem tökéletes képeket tudnak előállítani. Hogyan lehet megakadályozni, hogy ezeket megtévesztés, félrevezetés céljára használják?

– A Tengr.ai modellünk esetében már tettünk is bizonyos technikai lépéseket annak érdekében, hogy kevésbé legyen használható ilyen célokra. Ami biztosan lehet már szűrni a gépgeneráló szoftverek esetében, hogy ne gyártson megtévesztő képeket ismert személyiségekkel, közismert arcokkal. Ez a nagy amerikai startupok esetében is nagy kihívás volt, gondoljunk csak azon a kamufotók terjedésére, amelyek Donald Trump állítólagos letartóztatását ábrázolták.

Ugyanakkor mi nem cenzúrázunk bizonyos szavakat, ahogyan például a Midjourney alkalmazás teszi. Egyéb eszközeink vannak arra, hogy megakadályozzuk vagy legalább korlátozzuk a hamisítványok terjedését. Például egy ismert személyiség arca nem lesz „elég jó” ahhoz, hogy hiteles és meggyőző legyen, fel lehessen használni.

– Meg lehet-e egyértelműen különböztetni egy valódi fotót egy generált fotótól?

– A képgeneráló programok fotói már annyira közel állnak a valósághoz, hogy nehéz megkülönböztetni, hogy melyik a valódi, és melyik a mesterséges. Lehet persze félrevezető képet generálni, akár egy nem létező eseményről, de úgy gondolom, hogy ez nem csak a mesterséges intelligencia korára jellemző, hi-

szen a Photoshop is ugyanezt meg tudja csinálni, csak több időbe és munkába kerül. Nekünk nincsenek eszközeink arra, hogy megakadályozzuk az átverős képek terjedését a közösségi médiában: ezt a platformok kellene megtegyék, például a Facebook a fact-checking módszerrel.

– Hogyan képzeled el a fotózás jövőjét?

– Természetesen nem sportszerű, ha ugyanúgy ítéljük meg a mesterséges intelligencia által készült képet és a fotógéppel készített fotót. Ez nem tesz jót a képgeneráló programokkal foglalkozó szakembereknek sem, hiszen olyan, mintha csalásra adnának lehetőséget. Persze sokkal nehezebb lefotózni egy gepárdot a valóságban, valódi környezetében, mint promptolni, azaz szöveges utasítással elkészíteni az otthoni fotelből. Ez igaz bármely képre, amit a számítógéptől rendelünk, azzal együtt, hogy a mesterséges intelligencia esetében sokkal több kontrollom van a fotó felett, mint valós helyzetben: nem kell például a modellnek fizetnem, nem kell fényeket beállítanom a helyszínen, nem kell váltogassam a díszletet. Mindez könnyedén és gyorsan megoldható azokkal az utasításokkal, amiket a programnak adok, hogy milyen képet készítsen.

Én ahhoz hasonlítanám a jelenlegi helyzetet, mint amikor a 19. században megjelent a fotográfia, és versenyre kelt a portréfestőkkel, akiknek már megvolt egy jól behatárolt piacuk. Amint megjelent a fényképezés, a festők válságba kerültek, ugyanakkor egy részük ráérzett arra, hogy a fotógép ezt jobban megcsinálja, és átképezték magukat fotósnak. A fényképezés nem jelentette a festészet mint művészet eltűnését, sőt sok tekintetben kihívás elé állította a festőket, új irányzatoknak nyitotta meg az utat, és így jelentek meg a 20. századi ismert irányzatok, a kubizmus, dadaizmus, fauvizmus. Úgy gondolom, hogy a fotóművészetnek is jót fog tenni a mesterséges intelligencia megjelenése: manapság már számos fotós ötvözi a fotózást a mesterséges intelligencia lehetőségeivel. Egy másik pozitív hatás az lehet, hogy a fotózásnak újra kell értelmeznie magát a művészetek sorában, ezért áttörés lehet a következő évtizedekben a fotózásban is.

– Mennyire vagy optimista a mesterséges intelligencia hatásával kapcsolatban, milyen társadalmat fog eredményezni az alkalmazása?

– A távolabbi jövőre nézve mindenképp optimista vagyok. Úgy, ahogy a korábbi technikai forradalmak esetében történt, a jelenlegi alkalmazások is rengeteg pozitív dolgot fognak eredményezni, könnyebb lesz a következő generációk élete, hatékonyabban fognak dolgozni a mesterséges intelligencia miatt. Ugyanúgy, ahogy a múlt század végén elterjedt az ötnapos munkahét, a következő években például valószínűleg el fog terjedni a négynapos munkahét, amit már jelenleg is alkalmaz több számítástechnikai cég. Ezt azért lehet megtenni, mert a mesterséges intelligencia a programozásban is szerepet játszik, sokkal gyorsabban lehet elkészíteni a fejlesztéseket. El fogunk jutni oda, hogy a szoftverfejlesztés is promptolás lesz: ugyanazt a szoftvert kevesebb fejlesztő fogja gyorsabban kidolgozni. A mesterséges intelligencia hatalmas hozzáadott értéket fog teremteni, ami társadalmi szinten is meg fog mutatkozni a társadalmai jólétben.

Sajnos rövid távon nem vagyok ennyire optimista. A társadalom át fog alakulni, munkahelyek fognak megszűnni, sok munkavállaló fog arra kényszerülni, hogy átképezze magát. Ugyanakkor lesznek emberek, akik arra fogják használni az új technológiát, hogy ártsanak. Gondolok itt a már említett álhírek terjesztésére, az emberek félrevezetésére, de ennél durvább dolgokat is el lehet képzelni: léteznek olyan drónok, amelyek már arcfelismerés segítségével képesek emberekre támadni, és ezek a technológiák csak fejlődni fognak a jövőben.

– Az egyik korábbi interjúdban nyilatkoztad, hogy a mesterséges intelligencia a demokrácia halálát jelentheti. Mennyire reális ez a veszély szerinted?

– Én úgy gondolom, hogy a demokráciának része a párbeszéd, az egyenlő esély arra, hogy az üzenetek eljussanak a választókhoz. A jelenlegi közösségi médiában rengeteg olyan eszközt használnak, amelyek a mesterséges intelligenciához kapcsolódnak. Ha ezeket az eszközöket egy bizonyos hatalom uralja, akkor a visszhangkamra-effektus miatt egyes üzenetek felerősödnek, az ellentétes üzenetek meg egyáltalán nem jutnak el a közönséghez. Például az orosz–ukrán háború kapcsán látható, mit jelent a híráramlás tekintetében, hogy Oroszországnak van egy saját közösségimédia-oldala, a VKontake, amin keresztül szűrt információt kap az orosz közönség. De arra is lehet hivatkozni, hogy Oroszországban nem a Google-keresőt használják, hanem van saját keresőmotorjuk, a Yandex, ami szintén az információ szűrésének az eszköze lehet.

Ugyanakkor az Európai Unió szempontjából is fenyegetésnek látom, hogy az információáramlás ki van szolgáltatva amerikai nagyvállalatoknak, amelyek nem tartják kötelezőnek magukra vonatkozólag az uniós jogszabályokat. Jelenleg az amerikai fejlesztésű ChatGPT-nek akkor túlsúlya van a piacon, hogy sokan nem is hallottak egyéb modellekről, holott egész sor más modell létezik már. A megoldás az lenne, hogy az Európai Uniónak is legyenek ütőképes modelljei, hiszen egyáltalán nem mindegy, mire tanítják a modelleket; a mesterséges intelligencia területe lehet akár a következő hidegháború helyszíne.

– A társadalom mennyire van felkészülve ekkora szerkezeti váltásra?

– Szerintem biztos, hogy a mi generációnk, a jelenlegi harmincas, negyvenes korcsoport sincs erre felkészülve. Viszont abban reménykedem, hogy a következő generáció, a fiaink és unokáink már élvezni fogják a mesterséges intelligencia pozitív hozadékait. A következő évtizedek olyan mértékű társadalmi változásokat és átalakulásokat fognak eredményezni, amit most még nem is tudunk felmérni. Ahhoz tudnám hasonlítani, amikor az iparosodás következtében az agrártársadalom átalakult, gyári munkások vették át a földmunkások helyét. Két teljes generációnyi időre volt szükség, amíg kialakult egy új világrend; ez nem volt fájdalommentes, nagy törés volt az emberiség történetében. De hosszú távon pozitív hatásai voltak, ma már nem tudnánk elképzelni az életünket villanyáram és gépkocsi nélkül.

Még mi sem, akik mesterséges intelligenciát fejlesztünk, nem vagyunk teljesen felkészülve a jövőre, nem beszélve az emberiség nagy részéről. Remélem, hogy nem lesz hirtelen ugrás, hanem egy emberséges folyamat, aminek hatására fokozatosan egyre több gépi tanuláson alapuló eszközt használunk. A folyamat már elkezdődött, említhetném akár az okostelefonokat, ahol már rengeteg olyan megoldás fut, aminek gépi tanulás az alapja.

Kérdezett Győrffy Gábor