

JAK UMĚLÁ INTELIGENCE MYSLÍ

JAK FUNGUJE GENERATIVNÍ UMĚLÁ INTELIGENCE

Jak funguje generativní umělá inteligence

Generativní umělá inteligence je typem umělé inteligence, která je schopna vytvářet obsah v různých formách, jako jsou například nové texty, obrázky, videa, hudba a další. Tento obsah vytváří na základě požadavků (příkazů) od uživatele neboli na základě promptů. Modely generativní umělé inteligence v zásadě fungují tak, že v reakci na náš prompt a na základě dat, ze kterých se učily, vygenerují tu nejvíce pravděpodobnou odpověď. Příkladem takového promptu může být: „Napiš pět hlavních důvodů, které vedly k rozdělení České a Slovenské Federativní Republiky v roce 1993.“

Jak umělá inteligence reaguje na text

Aby byla generativní umělá inteligence, jako jsou například modely GPT nebo Gemini, schopna odpovědět na prompt a vygenerovala např. požadovaný text, musí se „natrénovat“ na velkém množství již existujících dat. Díky tomu je pak schopna generovat nový obsah. Proto je tedy potřeba naučit model umělé inteligence rozpoznávat vzory a struktury na trénovacích datech, které danému modelu lidé poskytnou. Na základě těchto vzorů se pak model učí tvořit nová data, která jsou podobná těm, na kterých se učil. Pokud např. chceme, aby model byl schopen rozpoznat obrázky koček, je třeba mu poskytnout velké množství různorodých fotek koček. Model se na základě těchto fotek naučí rozeznávat vzory toho, jak vypadá kočka.

Z čeho se umělá inteligence učí

Trénovací data mohou být jakákoliv: můžeme použít databázi fotek nebo třeba zápisy ze soudních přelíčení. Pokud například chceme, aby model uměl dobře generovat text, dává smysl jako vstupní data použít velké množství psaného textu, který má určitou záruku kvality. V příkladu generování textu je možné jako vstupní data použít např. články velkých mediálních domů, které disponují velkým počtem psaných textů, ale také jsou jistou garancí kvality.

Kvalitu dat by pak měl zajistit proces čištění, kdy jsou trénovací data zejména strojově, ale i lidmi upravována tak, aby měla co největší kvalitu a neobsahovala nežádoucí informace. Je důležité mít vždy na paměti, že kvalita modelu umělé inteligence se odvíjí od dat, na kterých se daný model učí. Jinými slovy: umělá inteligence převezme to, co se naučí z dat. V tuto chvíli existuje velké množství relativně kvalitních dat, ale jejich zpracování nutné ke správnému fungování generativní umělé inteligence je velmi náročné. I proto si vývoj generativní umělé inteligence mohou finančně dovolit především jen velké technologické firmy, jako jsou Google a Microsoft.

Jak umělá inteligence generuje výstupy

Jakmile se model generativní umělé inteligence naučí na základě trénovacích dat reagovat na to, co po něm v zadaném promptu požadujeme, je schopen vygenerovat odpověď. Ta vzniká pomocí dalšího velmi podstatného principu fungování umělé inteligence – pravděpodobnosti. Jedná se totiž o statistické modely. Z toho také vyplývá její největší omezení.

Kdybychom ji například poskytli velké množství dat, kde se říká, že hlavním důvodem rozpadu České a Slovenské Federativní Republiky byli Martané, bylo by toto tvrzení součástí odpovědí, kterou by nám umělá inteligence velmi pravděpodobně vygenerovala.

Pokud například chceme textový výstup, model předpovídá pravděpodobnost každého dalšího slova v řadě na základě textu, na

kterém byl model vycvičen a na základě zadaného promptu. Modely umělé inteligence tak de facto nerozumí našim promptům nebo obsahu textů, ze kterých se učí, ale pouze vyhodnocují pravděpodobnost reakce. To může vést k vygenerování obsahu, který není pravdivý, protože umělá inteligence pouze vygeneruje výstup, který považuje za nejvíce pravděpodobný. To ovšem neznamená, že nutně odpovídá realitě.

Silné stránky generativní umělé inteligence

Jednou z klíčových výhod generativní umělé inteligence je její schopnost vytvářet nové a jedinečné výstupy. Například model umělé inteligence, který bychom vycvičili na velkém množství dat o probíhající klimatické změně a jejích příčinách, by byl schopen přijít s novými myšlenkami, jak jí čelit. Obdobně model vycvičený na datech s obrázky koček bude schopen namalovat obrázky, které dosud nikdo nevytvořil. Modely umělé inteligence mohou také generovat příběhy, postavy, a dokonce i celé dějové linie. Tyto modely dokáží analyzovat obrovské množství existující literatury a médií, naučit se vzorce a struktury vyprávění a poté generovat nová vyprávění, která se těchto principů drží a zároveň přinášejí nové a nečekané prvky.

Limity a slabé stránky

Potenciální nesprávnost výstupů je pravděpodobně hlavní nevýhodou pro využití umělé inteligence. Dá se říct, že generativní umělá inteligence se obvykle lépe hodí pro tvorbu obsahu založeného na kreativitě a interpretaci než na generování tvrdých a objektivních fakt. Dalším negativem je možné převzetí tzv. biasu (předpojatosti) z trénovacích dat, což může vést k dalším typům nežádoucích výstupů. Lidé by měli mít na paměti, že vygenerované výstupy mohou obsahovat např. stereotypní vyobrazení společenských menšin. V současnosti jsou také etickým problémem zdroje trénovacích dat, které jsou často získávány bez souhlasu vlastníků těchto dat. Je tedy otázkou, jestli to není v rozporu s autorskými právy. V neposlední řadě je problematický fakt, že umělá inteligence usnadňuje vytváření falešného nebo deepfake obsahu, což může vést k sofistikovanějšímu a snadnějšímu šíření dezinformací.