

Na FAV zkoušejí umělou inteligenci z vysokoškolské matematiky a logiky

29.7.2026 - Petr Girg, Martina Batková | Západočeská univerzita v Plzni

Nakolik si současné velké jazykové modely dokáží poradit se složitými matematickými úlohami, zjišťoval tým z Fakulty aplikovaných věd ZČU. Vybrané modely otestoval na osmnácti úlohách. Nejlepší výsledky nabídly největší modely, jejich postup ale nebyl vždy správný. Úspěšnost navíc závisí i na oboru.

Jaké jsou limity současné umělé inteligence v akademickém prostředí? Odpověď hledali studující z Fakulty aplikovaných věd Západočeské univerzity v Plzni (FAV). Vyvinuli systém, který automaticky zadává příklady velkým jazykovým modelům, zaznamenává výsledky i s postupy řešení a tím umožňuje průběžně vyhodnocovat jejich úspěšnost. Tým složený ze studentů i akademiků otestoval čtyři různé jazykové modely na osmnácti úlohách. Předběžná data z první fáze testování už ukazují jasné trendy.

Z více než sto dvaceti hodnocených odpovědí vyplynulo, že největší modely dosahují 98 procent úspěšnosti, zatímco menší varianty uspějí zhruba v 68 procentech případů. Podle vedoucího studentské části týmu Kristiana Tichoty je ale rozdíl částečně výsledkem technických omezení, malým modelům totiž chyběl přístup k externím nástrojům nebo narážely na limit délky textu.

Testy navíc potvrdily, že správný výsledek automaticky neznamená bezchybný postup. *„Zcela správnou logiku mělo jen 50 až 75 % odpovědí od velkých modelů a skutečně efektivní, přímočaré řešení nabídla nanejvýš polovina všech testovaných výstupů,”* řekl Tichota. Poukázal také na rozdíl v úspěšnosti řešení mezi jednotlivými matematickými disciplínami. V některých z nich už podle mentora projektu Petra Girga z katedry matematiky FAV současné velké jazykové modely dosahují úrovně vyučujících.

Vývoj a provoz testovací platformy propojil studující z kateder informatiky, kybernetiky a matematiky. Odborné vyhodnocení výsledků včetně kontroly postupů zajišťují akademici. Spolupráce napříč obory podle nich jasně ukázala, jak efektivně lze skloubit softwarové inženýrství s matematikou. *„Výsledný systém nám vyučujícím kontinuálně poskytuje přesná data o aktuálních limitech umělé inteligence v akademickém prostředí a umožňuje sledovat vývoj výkonnosti velkých jazykových modelů v čase. Pro studenty navíc představuje cennou praxi,”* řekl Petr Girg. Výsledky už tým prezentoval na studentských konferencích v Česku i zahraničí.

<https://info.zcu.cz/Na-FAV-zkouseji-umelou-inteligenci-z-vysokoskolske-matematiky-a-logiky/clanek.jsp?id=9845>