

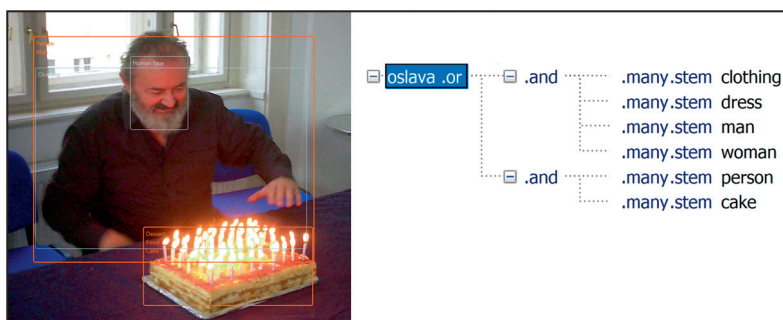
# Umělou inteligenci je potřeba zapojit do rozhodovacích procesů přirozeným způsobem

MIROSLAV NEČAS

Když jsem se před více než patnácti lety setkal na univerzitě s umělou inteligencí, bylo to v předmětu ekonomicko-matematické metody.

Tehdy jsme neuronovou síť tvořili pomocí maker v Excelu. Technologie se od té doby hodně posunula a strojové učení je dnes jednou z běžně používaných metod zpracování dat. Pořád jde ale o metodu stochastickou. Hodí se všude tam, kde řešíme opakované slabě strukturované problémy a spíše než o optimální řešení jde o řešení dostatečné. V našich řešeních aplikujeme strojové učení na rozpoznání entit v textu, objektů v obraze nebo převod mluveného slova do textu. To jsou úlohy, které nelze z kapacitních důvodů dělat ručně, a deterministické algoritmy se na ně nehodí.

Pokud extrahujeme entity z textu (abychom věděli, že se hovoří o určité osobě, firmě či místě), u entity typu e-mailu funguje deterministický algoritmus natolik spolehlivě, že strojové učení vůbec nemá smysl použít. Při extrakci entity typu telefonní číslo nebo osoba je stále deterministický algoritmus spolehlivější než strojové učení. Zejména díky tomu, že se můžeme opřít o znalost různých formátů telefonních čísel nebo znalost všech jmen a příjmení uživatelů v ČR. Avšak u entit typu firmy či adresy již strojové učení dosahuje lepších výsledků. Je to tím, že jména firem mohou být obecná slova, rovněž je potřeba podle kontextu rozlišit mezi názvem organizace a jménem výrobku. S adresami je to podobné. Řada ulic a míst se jmenuje po významných osobnostech atd. Při detekci objektů v obraze či při převodu hlasu do textu jsou výsledky strojového učení dnes již spolehlivější než výsledky deterministických algoritmů.



Při využití strojového učení je třeba stále mít na paměti několik věcí. Jde o stochastický algoritmus, proto je výsledek vždy do určité míry náhodný. Je daná kvalitou modelu a ta je daná kvalitou trénovacích dat. Jsou-li v trénovacích datech chyby, bude systém dávat chybné výsledky. Není-li trénovací množina dat úplná, systém si s některými problémy neporadí. Výsledek strojového učení je zkrátka konzistentní jen tehdy, když se zásadně neliší parametry provozních dat a parametry dat, na kterých byl systém natrénován. Změní-li se významně vstupní data, systém přestane dávat uspokojivé výsledky. Nepříjemné je, že u strojového učení automaticky nevíme, které parametry dat jsou pro určitý model významné.

Těchto omezení jsme si v Toveku dobře vědomi. Vyvinuli jsme proto vlastní postupy a nástroje pro přípravu trénovacích dat a ověřování natrénovaných modelů. To nám umožňuje zrychlit a zlevnit práci spojenou s manuálním značkováním dat, eliminovat chyby pramenící z lidské nepozornosti a zároveň rozpoznat, které parametry dat jsou pro určitý model významné. Například při tagování obrázku přímo definujeme oblast obrazu, jež je pro

Výsledek vyhledávání „Oslava“ na základě definovaného dotazu

**Místo nahrazení lidí stroji se snažíme ulehčit jim od rutinních úloh a vytvořit prostor pro práci s vysokou přidanou hodnotou.**

rozpoznání objektu určující. Jinak by se třeba mohlo stát, že se systém naučí identifikovat různé objekty podle doprovodného textu, nikoliv podle toho, jak vypadají.

Využití umělé inteligence v našich řešeních umožňuje významné rozšíření různorodosti dat, která jsme schopni využívat pro vyhledávání informací, a rovněž významné rozšíření způsobu vyhledávání údajů. Příkladem je prohledávání fotek, které zachycují konkrétní situaci, např. oslavu. Tak jako v textu vyhledáváme informace podle struktury klíčových slov, ve fotkách vyhledáváme podle struktury rozpoznávaných objektů, např. více než pět osob, láhev vína a dort. Tam, kde hledanou informaci nelze popsat, můžeme zase vytvořit model pro její rozpoznání, např. podpisová doložka v dokumentu nebo vyjádření názoru na určité téma. Tyto možnosti jsou velmi důležité pro využití obsahu stále rostoucího objemu fotek, videí a hlasových záznamů, pro vyšetřování incidentů či podvodů, různé kontroly a tvorbu znalostí o lokalitách, produktech nebo tématu.

Samozřejmě existuje řada úloh, na které se umělá inteligence vložně nehodí. Kromě dobře strukturovaných problémů, na něž se více hodí deterministické algoritmy, jsou to i problémy, které se neopakují tak často, aby bylo na čem systém učil. Kapitolou samou pro sebe jsou situace, které vyžadují kreativní přístup a improvizaci. Na ty se žádný systém nehodí. Proto potřebujeme a budeme potřebovat stále více chytré a kreativní lidi. Místo toho, abychom usilovali o nahrazení šikovných lidí stroji, snažíme se ulehčit jim od rutinních a opakovaných úloh a poskytnout jim více času na práci s vysokou přidanou hodnotou. Proto kombinujeme strojové učení, deterministické algoritmy a vizualizaci informací, tak abychom do rozhodovacího procesu dokázali efektivně zapojit umělou inteligenci i lidský faktor. Věříme, že jedině s takovými nástroji budou naši klienti dlouhodobě konkurenceschopní. Budou pracovat efektivně a zároveň se dovedou adaptovat na měnící se podmínky na trhu.

Autor je business development manager, TOVEK.