

Strojové učení v klinické medicíně

Umělá inteligence je obor informatiky zabývající se využitím výpočetní techniky pro řešení komplexních úloh z praxe, zpracování obrovských objemů dat, zpracování přirozeného jazyka pomocí počítačů a podobně.

Strojové učení je oblastí umělé inteligence, která využívá zejména statistické nástroje a umožňuje počítačům *učit* se opakující se vzorce z dřívějších dat. Díky tomu je strojové učení velmi dobře využitelné pro klasifikaci (tj. *onálepkování dat* pomocí klasifikace, např. zhoubný nádor – nezhoubný nádor), regresi (tj. předpověď číselných hodnot) a clustering (česky asi nejlépe *shlukování*, rozdělení datové sady, např. pacientů, na skupiny s podobnými vlastnostmi).

Úvod do problematiky

Umělá inteligence (*artificial intelligence, AI*) a strojové učení (*machine learning, ML*) jsou obory přinášející zásadní změny a pokrok ve zpracování dat do spousty oblastí lidské činnosti. Zdravotnictví rozhodně není výjimkou. Kliničtí lékaři se v blízké budoucnosti ocitnou v situaci, kdy budou moci využívat informační sítě schopné prozkoumat mnohem komplexnější souvislosti, než umožňuje lidský mozek. Proto bude výhodou, bude-li lékař znát principy umělé inteligence a strojového učení. Stejně jako kdysi se lékaři učili používat zobrazovací metody jako rentgen či ultrazvuk, aby *viděli* to, co lidský zrak neumožňuje spatřit, budou se nyní učit, jak používat strojové učení pro analýzu dat a klinické rozhodování v situacích, které mozek rozumem zhodnotit nedokáže.

Lidé odedávna sbírali různá data, astronomové sledovali pohyb nebeských těles, přírodovědci sledovali životní cykly organismů, fyzikové zapisovali spousty různých měření. Pro zpracování těchto dat nejprve používali vlastní úsudek a rozum, v průběhu staletí se do procesu zapojovala statistika. Od 17. století statistika zaujímala stále významnější místo ve vědě. Po nástupu počítačů již nebylo nutné spousty výpočtů provádět ručně. Pro analýzu bylo totiž možné využít výpočetní výkon strojů.

Nyní se vlastně dostáváme k tomu, co je v současnosti strojové učení, potažmo umělá inteligence. Vlastně se jedná o pokročilé využití statistiky. Pojďme si to ukázat na velmi jednoduchém příkladu:

Dejme tomu, že máme diabetologa a ten nám ukáže některé své nové pacienty před začátkem léčby:

Lačná glykémie (mmol/l)	Diabetik
3,4	ne
5,3	ne
7,1	ano
6,9	ne
8,9	ano

Nemusíme být lékaři a už od pohledu můžeme odhadnout, že hranice lačné glykémie pro diagnostiku diabetu se pohybuje mezi 6,9 a 7,1 mmol/l. Můžeme si udělat statistiku a zjistíme, kolik procent diabetiků má lačnou glykémii nad nějakou hodnotou. Strojové učení tohle všechno dokáže spočítat samo (pomocí statistiky) a po příchodu nového pacienta do ambulance praktického lékaře dokáže říct, zda je daný člověk diabetik nebo ne podle ranní lačné glykémie.

Strojové učení tedy potřebuje předchozí data, na základě kterých pak při sběru nových dat odhadne, do jaké skupiny nová data patří.

Ve výše uvedeném případě nám strojové učení moc nepomůže, že? Jednak kritéria diabetu známe a jednak máme po ruce diabetologa. Ale představte si, že máte za úkol:

- analyzovat 50 popisných hodnot histologického vzorku u 100 tisíc pacientů, jejichž cílem bude diagnostikovat nebo vyloučit zhoubný nádor,
- analyzovat 150 hodnot týkajících se životního stylu u 10 milionů pacientů, jejichž cílem bude vytvořit vzorec pro odhad rizika rozvoje kardiovaskulárního onemocnění s označením nejzásadnějších rizikových faktorů,
- vytvořit klinický systém, který bude varovat lékaře při zvýšení rizika rozvoje pooperační sepse na odděleních ARO a JIP nad 1 % na základě laboratorních odběrů u pacientů v kritickém stavu,
- vytvořit program, který analyzuje rentgenový snímek plic a upozorní lékaře, je-li tam přítomno ložiskové zastínění a zároveň navrhne další diagnostický postup.

Zdají se vám tyto scénáře příliš složité? Nejsou. Strojové učení pracuje na principu statistiky a pokud dodáme algoritmu dostatek vstupních dat, dokáže nám statisticky zhodnotit i nová data.

Například v Indii nyní funguje pilotní program, který dokáže zhodnotit sken sítnice, ukázat statistickou pravděpodobnost různých diagnóz a doporučit léčebný postup. V zemi, kde je fatální nedostatek očních lékařů a čekací doby jsou nekonečné, je takový systém doslova zrak zachraňující.

Zcela jistě jste se již setkali s výsledky strojového učení:

- naučil-li se váš mobil rozpoznat vaši tvář, aby vám odemkl obrazovku (*zpracování obrazu*),
- používáte-li e-mailovou schránku se spamovým filtrem (*klasifikace*),

- přeložili-li jste si někdy nějaký cizojazyčný text pomocí webového překladače (*zpracování přirozeného jazyka*),
- pokud jste se dívali na předpovědi volebních výsledků (*regrese*),
- jestliže jste nakupovali na e-shopu a viděli pod košíkem nabídku se souvisejícími produkty, které *lidé jako vy obvykle kupují* (*clustering*),
- a podobně.

Strojové učení je oborem umělé inteligence, počítačových věd, jejichž cílem je pomocí analýzy dat nalézt vzorce aplikovatelné na nová data. Nejde tedy o systémy, které by měly emoce, uměly samy něco vynalézat nebo vymýšlet. Jejich hlavní schopností je schopnost učit se a nalézat opakující se vzorce v datech. Tak například mohou počítače hrát šachy pouhou analýzou milionů šachových partií, aniž by do nich nějaký programátor vložil pravidla hry. Počítač zkrátka *vidí*, jak jednotlivé figurky táhnou a *zná* také postup při různých situacích, které ve hře mohou nastat.

Stejně to bude i v klinické medicíně. Programy se strojovým učením *uvidí*, jak jednotlivé nemoci probíhají, a dokáží navrhnout pravděpodobnou diagnózu a pravděpodobně ideální diagnosticko-terapeutický postup. Definitivní rozhodnutí ale vždy zůstane na člověku.

Základní druhy úloh

Vytváření modelů strojového učení v klinické medicíně

Záludnosti, které nás mohou potkat

Současnost a vývoj do budoucna

Odkazy

Externí odkazy

Použitá literatura

- ROWE, Michael. An Introduction to Machine Learning for Clinicians. *Academic Medicine*. 2019, roč. ?, vol. ?, s. 1, ISSN 1040-2446. DOI: 10.1097/acm.0000000000002792 (<http://dx.doi.org/10.1097%2Facm.0000000000002792>).
- LANTZ, Brett. *Machine Learning with R : Learn How to Use R to Apply Powerful Machine Learning Methods and Gain and Insight Into Real-world Applications*. - vydání. Packt Publishing, 2013. 396 s. ISBN 9781782162148.

Reference