

Úvod do strojového učení pro pathology

Tomáš Brázdil¹, Adam Kukučka¹, Vít Musil¹, Rudolf Nenutil², Petr Holub³

¹Fakulta informatiky, Masarykova univerzita, Brno, Česká republika

²Masarykův onkologický ústav, Brno, Česká republika

³Ústav výpočetní techniky, Masarykova univerzita, Brno, Česká republika

SOUHRN

Digitalizace postupně proniká do velké části medicínských oblastí včetně patologie. Společně s digitálním zpracováním dat přichází aplikace metod umělé inteligence za účelem zjednodušení rutinních procesů, zvýšení bezpečnosti apod. Ačkoliv se obecné povědomí o metodách umělé inteligence zvyšuje, stále není pravidlem, že by odborníci z netechnických oborů měli detailní představu o tom, jak takové systémy fungují a jak se učí. Cílem tohoto textu je přístupnou formou vysvětlit základy strojového učení s využitím příkladů a ilustrací z oblasti digitální patologie. Nejedná se samozřejmě o ucelený přehled ani o představení nejmodernějších metod. Držíme se spíše úplných základů a představujeme fundamentální myšlenky, které stojí za většinou učících systémů, s použitím nejjednodušších modelů. V textu se věnujeme zejména rozhodovacím stromům, jejichž funkce je snadno vysvětlitelná, a elementárním neuronovým sítím, které jsou hlavním modelem používaným v dnešní umělé inteligenci. Pokusíme se také popsat postup spolupráce mezi lékaři, kteří dodávají data, a informatiky, kteří s jejich pomocí vytvářejí učící systémy. Věříme, že tento text pomůže překlenout rozdíly mezi znalostmi lékařů a informatiků a tím přispěje k efektivnější mezioborové spolupráci.

Klíčová slova: Strojové učení – Digitální patologie – Neuronové sítě – Rozhodovací stromy – Diagnostické systémy

Introduction to Machine Learning for Pathologists

SUMMARY

Digitalization has gradually made its way into many areas of medicine, including pathology. Along with digital data processing comes the application of artificial intelligence methods to simplify routine processes, enhance safety, etc. Although general awareness of artificial intelligence methods is increasing, it is still not common for professionals from non-technical fields to have a detailed understanding of how such systems work and learn. This text aims to explain the basics of machine learning in an accessible way using examples and illustrations from digital pathology. This is not intended to be a comprehensive overview or an introduction to cutting-edge methods. Instead, we use the simplest models to focus on fundamental concepts behind most learning systems. The text concentrates on decision trees, whose functionality is easy to explain, and basic neural networks, the primary models used in today's artificial intelligence. We also attempt to describe the collaborative process between medical specialists, who provide the data, and computer scientists, who use this data to develop learning systems. This text will help bridge the knowledge gap between medical professionals and computer scientists, contributing to more effective interdisciplinary collaboration.

Keywords: Machine Learning – Digital Pathology – Neural Networks – Decision Trees – Diagnostic Systems

Cesk Patol 2025; 61(1): 11–21

V dnešní době proniká umělá inteligence do velké části oborů lidské činnosti. Patologie je jednou z oblastí, kde výzkum možných aplikací učících metod probíhá již několik let. Prominentní postavení mají v této oblasti metody pro analýzu obrazu, které lze využít pro automatizovanou analýzu mikroskopických snímků tkání. V tomto směru probíhá již řadu let mnoho výzkumných aktivit a bylo publikováno velké množství vědecké literatury (1). Dále vývoj metod pro zpracování nestrukturovaných textových dat, což je v současné době hnací motor vývoje umělé inteligence obecně, poskytuje metody pro analýzu lékařských zpráv a výsledky lze tedy například využít pro digitalizaci archivů a automatizované generování nových zpráv v průběhu analýzy vzorků (2).

V tomto textu prezentujeme základy strojového učení ilustrované na příkladech z digitální patologie. V závěru se pokusíme naznačit, jaký je aktuální stav učících systémů v digitální patologii, a kterým směrem se bude patrně ubírat další vývoj. Tento

článek je první ze dvou článků o strojovém učení v digitální patologii. Jeho cílem je uvést základní principy a metody strojového učení formou přístupnou pro odborníky v oblasti patologie. Neklade si tedy za cíl představit nejmodernější technologie, ale zvýšit povědomí o učících systémech obecně. Druhý článek bude více věnován detailům propojení strojového učení s patologií a výsledným diagnostickým systémům stále více pronikajícím do praxe. Text se věnuje primárně postupům strojového učení pro klinickou praxi. Strojové učení je samozřejmě využíváno širěji jako podpůrná metoda ve výzkumu a analýze dat pro vědecké účely. Tomuto tématu se nebudeme primárně věnovat, nicméně společný základ v podobě principů strojového učení a základního vyhodnocení modelů prezentován bude.

UMĚLÁ INTELIGENCE A STROJOVÉ UČENÍ

Pojem umělé inteligence se často zaměňuje za určitý typ metod strojového učení. Zejména za metody postavené na velkých jazykových modelech, tedy ChatGPT (3) apod. Ovšem pojem umělé inteligence je daleko širší. Nahlédneme-li do známé učebnice (4), nalezneme v ní v podstatě průřez velkou částí informatiky. Počínaje plánováním, přes učení a reprezentaci znalostí, až po robotiku. V tomto textu se zaměříme na onu část, která se zabývá strojovým učením, tedy tvorbou počítačových systémů, jejichž funkcionalita nemusí být celá „ručně“ programována, protože se jí učí samostatně z dat.

✉ Adresa pro korespondenci:

Doc. RNDr. Tomáš Brázdil, Ph.D.

Fakulta informatiky, Masarykova univerzita, Brno

Botanická 68a, Brno

tel.: 549 49 3922

e-mail: xbrazdil@fi.muni.cz