

Z Wikipedie, otevřené encyklopedie

ikona

Tento článek potřebuje úpravy.

Můžete Wikipedii pomoci tím, že ho [vylepšíte](#). Jak by měly články vypadat, popisují stránky [Vzhled a styl](#), [Encyklopedický styl](#) a [Odkazy](#).

---

Konkrétní problémy: *nedostatečně ozdrojováno; heslovité sekce převést na souvislejší text, případně i dokončit*

**Strojové učení** je podoblastí [umělé inteligence](#), zabývající se [algoritmy](#) a technikami, které umožňují počítačovému systému „učit se“. Učením v daném kontextu rozumíme takovou změnu vnitřního stavu systému, která zefektivní schopnost přizpůsobení se změnám okolního prostředí.

Strojové učení má široké uplatnění. Jeho techniky se využívají pro rozpoznávání obrazů (např. ve formě [pixelů](#)) či [akustických](#) (např. [rozpoznávání řeči](#)) nebo elektrických (např. [EKG](#), [EEG](#)) [signálů](#), dále ke [klasifikaci](#), [segmentaci](#) či [kompresi](#) dat, k předvídání vývoje časových řad (např. [burzovních indexů](#)), k [analýze psaného textu](#) či k [filtrování spamu](#). V lékařství slouží k diagnostice onemocnění a v řízení [pro podporu rozhodování](#).

## Základní rozdělení algoritmů učení

[Algoritmy](#) strojového učení lze podle způsobu učení rozdělit do následujících kategorií:

- [Učení s učitelem](#) („supervised learning“) – pro vstupní data je určen správný výstup (třída pro klasifikaci nebo hodnota pro regresi).
- [Učení bez učitele](#) („unsupervised learning“) – ke vstupním datům není známý výstup.
- Kombinace učení s učitelem a bez učitele („semi-supervised learning“) – část vstupních dat je se známým výstupem, ale další data, typicky větší, jsou bez něj.
- [Zpětnovazebné učení](#) („reinforcement learning“), též *učení posilováním*.

Podle způsobu zpracování lze algoritmy rozdělit na:

- dávkové: Všechna data se požadují před začátkem výpočtu.

- inkrementální: Dokážou se „přiučit“, tj. upravit model, pokud dostanou nová data, bez přepočítání celého modelu od začátku.

## Základní druhy úloh

- [Klasifikace](#) rozděluje data do dvou nebo několika tříd ([učení s učitelem](#))
- [Regrese](#) odhaduje číselné hodnoty výstupu podle vstupu ([učení s učitelem](#))
- [Shlukování](#) zařazuje objekty do skupin s podobnými vlastnostmi ([učení bez učitele](#))

Další typy úloh jsou:

- [Ranking](#) určuje pořadí datových bodů, výsledkem je částečné nebo úplné setřídění
- Učení strukturovaných dat. Výstupní neboli hledaná struktura může být například sekvence, strom, graf, matice...

Aplikace jsou např. učení [syntaktických stromů](#) ve [zpracování přirozeného jazyka](#), zarovnání několika [sekvencí](#) proteinů v [bioinformatice](#), převod [řeči na textový](#) řetězec, tj. na [sekvenci](#) znaků, hledání vhodné molekuly reprezentované jako [graf](#) v [chemoinformatice](#), výstup [zpracování obrazu](#) jako [matice](#)...

## Podoblasti strojového učení

- Používané modely:
  - [Rozhodovací stromy](#)
  - [Algoritmus k-nejbližších sousedů](#)
  - Podpůrné vektory, viz [Support vector machines](#)
  - [Lineární diskriminační analýza](#) ([en:Linear discriminant analysis](#))
  - [Kvadratická diskriminační analýza](#) ([en:Quadratic discriminant analysis](#))
  - [Neuronové sítě](#), [učení s učitelem](#), [učení bez učitele](#), [hluboké učení](#)
  - [Bayesovské sítě](#)
- Techniky pro kombinaci více modelů ([en:Meta learning](#))
  - [Boosting](#)
  - [Stacking](#)

- [Bootstrap \(Bagging\)](#)
- Testování přesnosti modelu:
  - [Křížová validace](#)
  - [Diskriminační analýza](#)
  - [Bootstrap \(Statistics\)](#)

## Terminologie

- Data, body, případy, měření
- Atributy, rysy, proměnné, fíčury/features
- [Druhy/typy atributů](#): kategoriální (např. „A“, „B“, „AB“ nebo „O“ pro krevní skupiny), ordinální (např. „velký“, „střední“ nebo „malý“), strukturované, hierarchické, binární, celočíselné (např. počet výskytů slova v emailu) anebo reálné (např. měření krevního tlaku)

## Software

[RapidMiner](#), [KNIME](#), [Weka](#), [ODM](#), [Shogun Toolbox](#), [Orange](#), [Apache Mahout](#) a [Scikit-Learn](#) jsou softwarové balíky, které obsahují různé algoritmy strojového učení.

Online: Microsoft Azure Machine Learning (Azure ML), TensorFlow (Google)

## Hardware

Optimální hardwarová implementace provádění výpočtů umožňuje značně urychlit nalezení výsledku. Používají se různé způsoby:

- [Tensor Processing Unit](#) (TPU) <sup>[1]</sup> od [Google](#) jsou specializované integrované obvody ve formě čipů implementovaných v zařízení TPU (An application-specific integrated circuit (ASIC)), který optimalizuje výpočetní výkon strojového učení.
- [Graphics Processing Unit](#) (GPU) je systém procesorů implementovaných na grafické kartě, původně určený především pro běh počítačových her, kladoucích velké nároky na výpočetní výkon. Na rozdíl od [CPU](#) majících k dispozici maximálně 8 nebo 16 výpočetních vláken, GPU jich může využívat až např. 2 048. Pro algoritmizaci učení se pak může užít tzv.

[CUDA](#) platforma, umožňující programovat v jazycích [C++](#) resp. [Fortran](#), která umožní díky velkému počtu dostupných výpočetních vláken vysoký stupeň [paralelizace](#) výpočtu, čímž zvýší výpočetní výkon strojového učení.

## Reference

1. [↑](#) Google developers: [Machine Learning Glossary](#)

## Související články

- [Data mining](#)

## Externí odkazy

- Obrázky, zvuky či videa k tématu [strojové učení](#) na Wikimedia Commons

Citováno z „[https://cs.wikipedia.org/w/index.php?title=Strojové\\_učení&oldid=25304418](https://cs.wikipedia.org/w/index.php?title=Strojové_učení&oldid=25304418)“

**Kategorie:**

- [Strojové učení](#)
- [Umělá inteligence](#)

**Skryté kategorie:**

- [Údržba:Články k úpravě](#)
- [Monitoring:Články s identifikátorem NKC](#)
- [Monitoring:Články s identifikátorem GND](#)
- [Monitoring:Články s identifikátorem LCCN](#)
- [Monitoring:Články s identifikátorem NDL](#)
- [Monitoring:Články s identifikátorem NLI](#)

Hledání

Speciální:Hledání

Hledat

Strojové učení

88 jazyků

[Přidat téma](#)