

Učení bez učitele (Unsupervised Learning)

Učení bez učitele je typ strojového učení, kde algoritmus pracuje s daty, která **nejsou označená** (nemají žádné labely/štítky). Model nedostává žádné „správné odpovědi“, ale snaží se sám najít vzorce, podobnosti nebo anomálie v surových datech.

Je to podobné, jako kdybyste dali dítěti hromadu různobarevných kostek různých tvarů. Dítě neví, jak se barvy jmenují, ale samo je začne třídit na hromádky podle barvy nebo podle toho, které jsou kulaté a které hranaté.

1. Hlavní úlohy učení bez učitele

Tento přístup se využívá především pro tři typy úkolů:

A. Shlukování (Clustering)

Algoritmus rozdělí data do skupin (**clusterů**) tak, aby si data uvnitř jedné skupiny byla co nejvíce podobná a data z různých skupin co nejvíce odlišná.

- **Příklady:** Segmentace zákazníků (skupiny s podobným nákupním chováním), seskupování podobných zpráv v Google News.
- *Typické algoritmy:* **K-Means**, DBSCAN, Hierarchické shlukování.

B. Redukce dimenzionality (Dimensionality Reduction)

Snížení počtu proměnných (sloupců) v datech při zachování co největšího množství informací. Pomáhá to vizualizovat složitá data nebo zrychlit trénování jiných modelů.

- **Příklady:** Komprese obrázků, vizualizace genetických dat ve 2D grafu.
- *Typické algoritmy:* **PCA** (Principal Component Analysis), t-SNE, Autoenkodéry.

C. Asociační pravidla (Association Rules)

Hledání pravidel, která popisují velké části vašich dat.

- **Příklad:** „Lidé, kteří si koupili pivo, si s pravděpodobností 80 % koupili také brambůrky“ (Market Basket Analysis).
- *Typické algoritmy:* Apriori, Eclat.

2. Detekce anomálií (Anomaly Detection)

Specifické využití, kde model identifikuje datové body, které se výrazně liší od zbytku („outliers“).

- **Využití:** Detekce podvodů s kreditními kartami (neobvyklá transakce), identifikace poruch v průmyslových senzorech.

3. Výhody a nevýhody

Výhody	Nevýhody
Nepotřebujete drahá, ručně označená data.	Výsledky mohou být těžko interpretovatelné (proč model vytvořil právě tyto skupiny?).
Odhaluje skryté vzorce, které lidé nemusí vidět.	Je obtížné objektivně změřit „přesnost“ modelu (neexistuje správná odpověď).
Ideální pro průzkumnou analýzu dat (EDA).	Výstupy často vyžadují experta, který jim dá význam.

4. Praktické využití

- **Bioinformatika:** Seskupování genů s podobnými funkcemi.
- **Kybernetická bezpečnost:** Detekce neobvyklého provozu v síti (možný útok).
- **Doporučovací systémy:** Hledání podobných uživatelů pro doporučení filmů (Netflix).

Související články

- [Podrobně o shlukování](#)
- [Učení s učitelem](#)
- [Autoenkodéry a neuronové sítě](#)

Tagy: *ai ml unsupervised_learning clustering pca anomálie*

From:

<https://www.serviceit.cz/> - IT ENCYKLOPEDIE

Permanent link:

https://www.serviceit.cz/doku.php?id=it:ml:unsupervised_learning

Last update: **2026/01/02 12:43**

