

8. Základní principy umělé inteligence

🕒 Created	@April 26, 2025 7:11 PM
🏷️ Tags	Done
👤 Kdo vypracoval	Robb

Intelligence = schopnost myslet, vytvářet nové myšlenky, adaptace a učení, chápat koncepty a řešit problémy

Umělá inteligence = stroje, které by normálně potřebovali lidskou inteligenci

formální definice = věda a tvorba inteligentních strojů (John McCarty)

Experimenty pro zkoumání umělé inteligence

Turingův test

Autor: **Alan Turing**

Testuje zda umělá inteligence vykazuje inteligentní chování.

=Komunikace s člověkem, který rozhoduje jestli komunikuje se strojem nebo člověkem.

Kritika: Test měří jak umí stroj napodobit lidské chování ne skutečnou inteligenci.

Čínský pokoj

Autor: **John Searl**

Zpochybňuje že simulace inteligentního chování znamená že rozumí nebo že má vědomí (protiargument **Turingovo testu**)

=Člověk sedí v místnosti a dostává čínské znaky, na které odpovídá podle předem daných pravidel (pomocí slovníku atd.)

Kritika: Někteří si myslí, že systém jako celek by mohl vykazovat pochopení

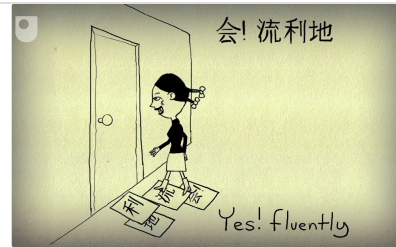
Krátké video o experimentu:

The Chinese Room - 60-Second Adventures in Thought (3/6)

An argument against computers ever being truly intelligent.

(Part 3 of 6)

 <https://www.youtube.com/watch?v=TryOC83PH1g>



Hard AI vs Soft AI

Soft AI (Narrow AI) = AI je zaměřena na konkrétní oblasti nebo úkoly, nemůže samostatně rozhodovat mimo definovaný rámec, nemá schopnost generalizace

- Rozpoznávání obrazu
- Hlasový asistent (Siri, Alexa, Google assistant) - schopni odpovídat a vykonávat úkoly podle hlasového rozkazu
- Doporučovací systémy (Spotify)
- Autonomní vozidla
- Prediktivní modely

Výhody:

- Vysoká efektivita a přesnost při řešení specifických problémů
- Snadná implementace
- Bezpečnější než Hard AI, protože nemůže přejít mimo svůj specifický úkol

Nevýhody:

- Neschopnost učit se a přizpůsobovat novým úkolům mimo oblast
- Závislost na datech a programování - omezená flexibilita
- Neumí porozumět širšímu kontextu nebo obecným problémům

Hard AI (General AI) = řešení obecných problémů, napodobení lidského uvažování, dokáže vyřešit širokou škálu problémů, přizpůsobí se novým situacím a řeší neznámé problémy

- Hypoteticky schopnost samostatného myšlení
- Prvky vědomí a kreativity
- Učí se a jedná nezávisle na lidech
- Neexistuje ZATÍM!!!!
- HAL 9000 (Vesmírná odyssea), Skynet, Knight Rider, Jarvis

Výhody:

- Potenciál řešit vysokou škálu problémů
- Schopnost přizpůsobit se novým podmínkám

Nevýhody:

- Technologicky náročné
- Etické otázky: Jak zajistit že nebude AI ohrožovat lidskou existenci
- Obava že se vysoce inteligentní systém začnou chovat nepředvídatelné
- Potřeba obrovských množství zdrojů

Nástroje umělé inteligence

Machine learning (ML)

Podobor AI, který umožňuje systému učit se na základě dat

Hlavní metody:

1. **Supervised learning** = Učení s učitelem (klasifikace, regrese)

Stroj učíme ručně = předložím obrázek slona a řeknu mu že je to slon 🤔

Zobecňování dat pomocí regrese a podle regrese třídí data


2. **Unsupervised learning** = Učení bez učitele (shlukování, redukce dimenzionality)

Určování závislosti a zjednodušování = redukce dimenzionality

shlukování

2.1. Gaussian mixture models

sklearn.mixture is a package which enables one to learn Gaussian Mixture Models (diagonal, spherical, tied and full covariance matrices supported), sample them, and estimate them from data. Facilit...

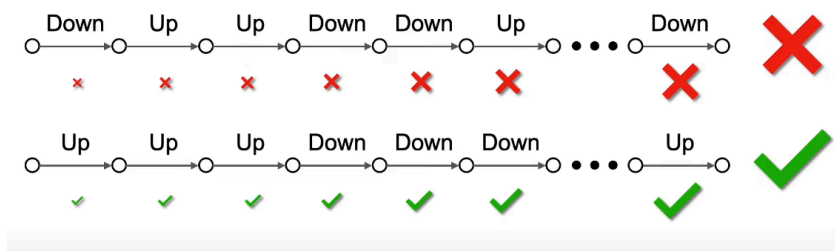
 <https://scikit-learn.org/1.5/modules/mixture.html>

3. **Reinforcement learning** = Posilované učení

Nehodnotíme jednotlivé kroky až výsledek

Neřešíme, které kroky vedly ke špatnému výsledku a všechny penalizujeme

Kroky můžeme penalizovat/odměnit podle toho kdy jsme krok provedli viz obrázek



Penalizace podle toho kdy byl krok vykonán

Na obrázku můžeme vidět dvě sekvence kroků ve hře pong (v sekvenci se posouvá odrážecí platforma). Horní sekvence skončila prohrou a dolní výhrou. Při ohodnocení kroků v horní sekvenci se kroky penalizovali podle toho, jak vzdálené konci jsou a tady podle toho jestli výsledkem sekvence bylo vítězství nebo porážka. Stejně tak byly odměněny body v dolní sekvenci.

Celé lépe popsáno v přiloženém videu

Př.: U piškvorek nebudeme hodnotit jednotlivé tahy ale až jestli prohrál nebo vyhrál

Reinforcement Learning from scratch

How does Reinforcement Learning work? A short cartoon that intuitively explains this amazing machine learning approach, and how it was used in AlphaGo and ChatGPT.

<https://www.youtube.com/watch?v=vXtfdGphr3c&t=171s>

Supervised learning

Regrese

Proložení dat funkcí (lineární, kvadratická...)

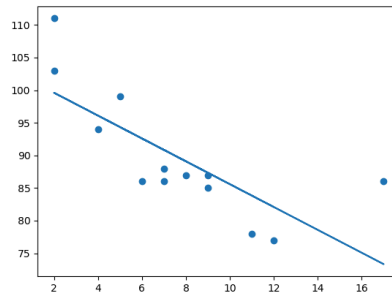
Chyby (epsilon) sčítáme umocněné na druhou (ze záporných to udělá kladné)

Přeučená a nedoučená soustava (mají méně nebo více neznámých než rovnic)

Nezávislá a závislá proměnná

Interpolace = nalezení takové funkce, která proleze všema bodama

Extrapolace = využití zkušeností z definovaného pole pro předpověď



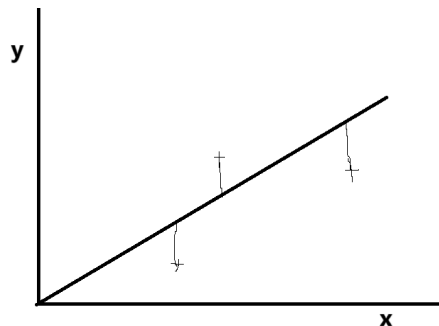
$$\sum_{i=1}^n (y - y_i)^2$$

Klasifikace

Třídění objektů

Lineární separace (sumisti a baletky)

Můžeme použít přímku (i více) pro separaci dat tříd



Nejbližší soused - problém když nejsou data tříd daleko od sebe

Data se třídí podle nejbližších **zástupců** (sousedů) tříd

K-NN (k nejbližších sousedů) - bereme v potaz k nejbližších sousedů

Můžeme zavést váhy sousedů podle vzdálenosti

Trénování a Testování

rozdělení trénovacích a testovacích dat např 9:1

data musíme náhodně zamíchat aby se nevytvořila systematická chyba

Unsupervised learning

K means

Vybereme 3 náhodné body (počet bodů záleží na předem určeném počtu skupin)
= inicializace

Do vybraných bodů se vloží středy clusterů

Podle vzdálenosti od středů clusterů se body zařazují do skupin

Poté se pro každý cluster spočítá střední hodnota (mean)

Spouští se několikrát

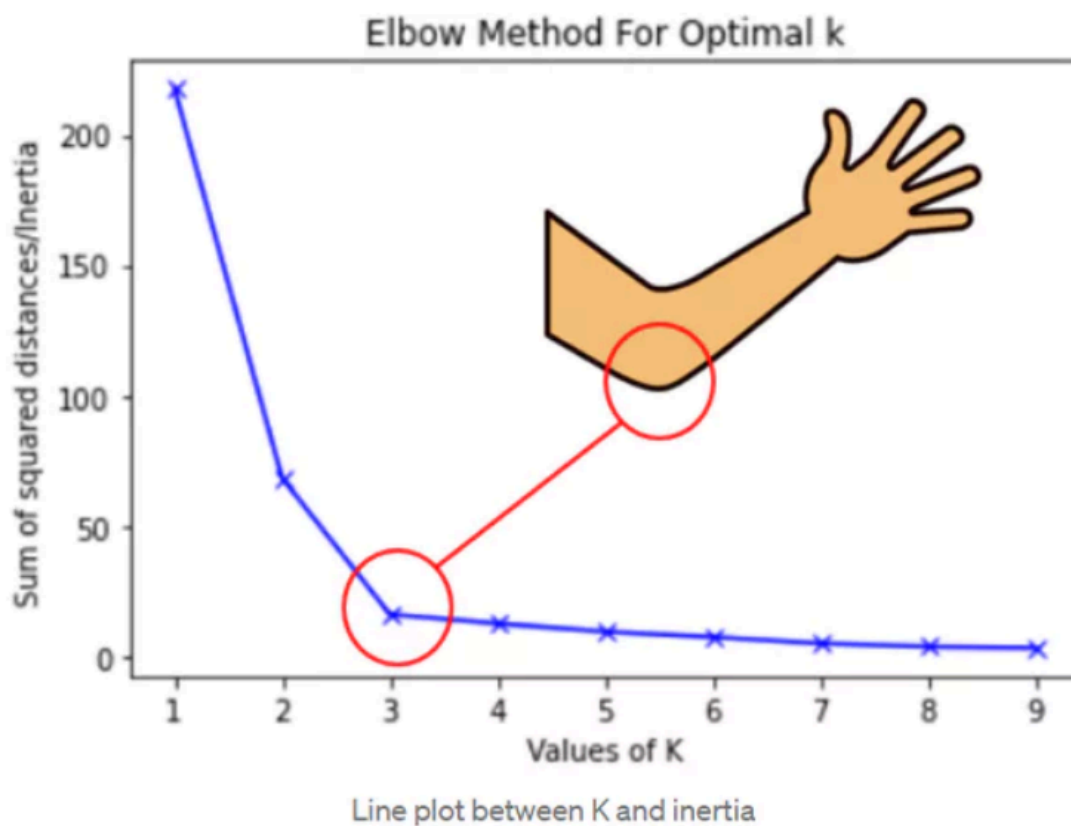
Počet clusterů určujeme dopředu!!!

Použití: V nákupním centru pomocí této metody určují jak slevit a zdražit aby měl zisk

Nevhodné pro promíchané data

Když chceme zjistit kolik shluků tam je:

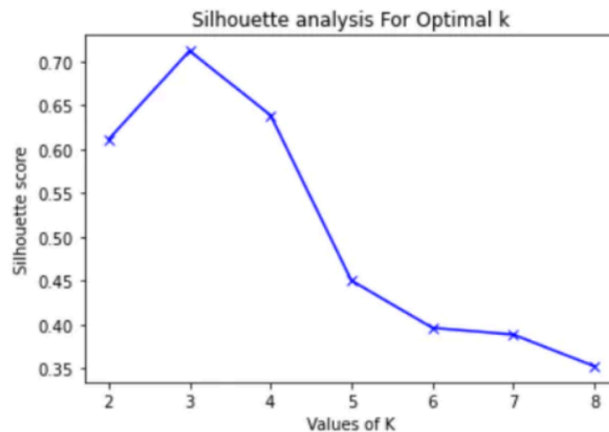
1. Zvolíme nejdříve 2 shluky
2. Vypočítáme chybu
3. Hledáme místo kde je "rameno"



Nebo také kdy je "Silhouette score" v maximu

$$S(i) = \frac{b(i) - a(i)}{\max \{a(i), b(i)\}}$$

Silhouette koeficient



Silhouette plotnuté, optimální počet clusterů je tedy 3

Redukce dimenzionality

Hledáme závislost výstupů na všech vstupech

Pokud zjistíme že na některých vstupech nezáleží nebo na nich málo záleží vyškrtneme je.

Reinforcement learning

Genetické algoritmy

Inspirované evoluční teorií a principem přirozeného výběru

Simulace evolučních procesů mutace, křížení a selekce pro optimální řešení

Silný nástroj pro hledání optimálních řešení (např.: minimalizace funkce)

Mutace = přidání drobné změny, která je dostatečně veliká aby se změnil výsledek

Inbreeding = Nový jedinec vychází ze kombinace nejlepších původních jedinců

Můžeme se stát že nalezneme jen lokální min/max → použijeme náhodnou mutaci

Také se může dočasně zhoršit výsledek ale díky němu se můžeme dostat přes hranici, která bránila zjistit skutečné min/max

Př.: když chceme najít minimum funkce

$$f_x = 3x^2 + 2x + 5$$

máme jedince:

P1 = 5→90

P2 = 3→38

P3 = -2→13

P4 = -5→70

P2.1 = -1→6

P2.2 = 2,1→22.43

P3.1 = -0,8→8,52

P3.2 = 0→5

...

Úloha obchodního cestujícího = hledání cesty mezi body, každá cesta je ohodnocena

Deep learning (DL)

Využívá hluboké neuronové sítě (Alespoň jednu neuronovou vrstvu mezi vstupní a výstupní) k porozumění dat

Typy neuronových sítí:

1. **Konvoluční neuronové sítě** (CNN) - zpracování obrazu
2. **Rekurentní neuronové sítě** (RNN) - zpracování sekvenčních dat
Využívá se u řeči protože řeč je sekvence dat za délku času
3. **Transformery** - zpracování přirozeného jazyka
Kombinace různých vrstev a jejich propojení (architektura)

Náročné na výkon

- Aplikace: rozpoznávání textu, autonomní vozidla, generace textu

Expertní systémy

Systémy založené na chování odborníků

Aplikace:

- Lékařské diagnózy
- Technická podpora
- Řízení rizik

Bayesovské systémy

Pravděpodobnostní model založen na Bayesově teorému. Ten umožňuje aktualizovat pravděpodobnost na základě nových dat

Pracují s nejistotou a umožňují efektivní rozhodování a predikce.

Hlavní metody:

1. **Naivní Bayes** = jednoduchý klasifikační algoritmus,
 2. **Bayesovské sítě** = grafické modely znázorňující pravděpodobnostní vztahy proměnných, diagnostika a modelování
 3. **Bayesovské interference** = postup aktualizace pravděpodobnosti podle nových dat
 4. **Bayesovské filtrování** = predikci a odstranění šumu v datech, GPS
-

Matouš

Robb