

Hluboké učení

28. listopadu 2023

10. Strojové učení

Hluboké učení (*Deep Learning*) je disciplína spadající do kategorie strojového učení. Zabývá se vývojem a aplikací algoritmů pro učení umělých neuronových sítí s hierarchickou reprezentací vlastností (angl. *features*). Učení může probíhat tzv. *s učitelem* (angl. *supervised*), *bez učitele* (*unsupervised*), nebo kombinací obojího (*semi-supervised*).

Hluboké učení umožňuje extrakci komplexních vlastností (angl. *features*) z neupravených dat. Například při zpracování obrazu nižší vrstvy modelu identifikují hrany v obrazových datech sousedících pixelů, zatímco vyšší vrstvy pracují s obecnějšími koncepty bližšími lidskému chápání, jako jsou číslice, písmena nebo tváře.

Mělké učení může dosáhnout podobných výsledků s méně prostředky, takže existuje i ekologičtější přístup.

Různé varianty hlubokého učení jsou využívány v oborech počítačového vidění, rozpoznávání řeči, zpracování přirozeného jazyka, strojového překladu, bioinformatiky, návrhu léčiv, analýzy lékařských snímků, výzkumu klimatu a hraní deskových her (kde tyto metody dosáhly srovnatelných nebo lepších výsledků než profesionální hráči).

10. Strojové učení

Definice

Pojem "hluboké" se vztahuje k počtu vrstev neuronů, ve kterých probíhá transformace dat. Přesněji řečeno jde o počet transformací dat cestou ze vstupu (vstupní vrstva) na výstup (výstupní vrstva). Pro dopředné neuronové sítě je hloubka sítě dána jako počet mezilehlých skrytých vrstev. Neexistuje proto obecně platná hranice, která by oddělovala hluboké učení od ne-hlubokého učení. Hloubka sítě musí být minimálně rovna 1, neboť síť už bude schopna aproximovat libovolnou vektorovou funkci. Přidávání dalších vrstev nezlepšuje aproximační schopnosti sítě, ale umožňuje extrahovat z dat komplexnější vlastnosti. Větší počet vrstev tak zefektivňuje proces učení. U rekurentních neuronových sítí může signál putovat každou vrstvou několikrát.^[12]

Pro odhad parametrů sítě (trénování) se obvykle používá algoritmus zpětného šíření chyby. Trénování probíhá ve dvou fázích, tj. nejprve předučení sítě dopředným směrem, např. pomocí autoenkodérů (učení bez učitele) a poté doučení sítě zpětným směrem (učení s učitelem). Eliminuje se tak tlumení zpětného šíření chyby.

10. Strojové učení

Revoluce hlubokého učení

V roce 2012 tým vedený Georgem E. Dahlem využil víceúlohovou hlubokou neuronovou síť k předpovědi biologického cíle konkrétního léku. V roce 2014 Hochreiterova skupina využila hluboké učení k detekci nežádoucích účinků chemikálií obsažených ve výživě, výrobcích pro domácnost a lécích na životní prostředí.

Další významné pokroky v oblasti rozpoznávání obrazu proběhly v letech 2011 a 2012. Rychlé konvoluční neuronové sítě využívající pro svůj běh GPU přinesly důležitý pokrok v oblasti počítačového vidění. V roce 2011 tento přístup poprvé vedl k překonání lidských schopností v rozpoznávání obrazových vzorů (patterns). Vedl rovněž k vítězství v soutěži rozpoznávání rukou psaného čínského písma. V říjnu roku 2012 podobný systém s převahou zvítězil v soutěži projektu ImageNet. V listopadu téhož roku další z těchto systémů zvítězil v ICPR soutěži v analýze lékařských snímků pro detekci rakoviny. V letech 2013 a 2014 se chybovost systémů v úloze ImageNet dále snižovala, podobný trend zaznamenalo rovněž odvětví rozpoznávání řeči.

10. Strojové učení

Dle některých výzkumníků bylo vítězství v ImageNet základem tzv. "revoluce hlubokého učení" (angl. "deep learning revolution"), která transformovala celé odvětví umělé inteligence.

10. Strojové učení

Dle některých výzkumníků bylo vítězství v ImageNet základem tzv. "revoluce hlubokého učení" (angl. "deep learning revolution"), která transformovala celé odvětví umělé inteligence.

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

- Tensor Processing Unit (TPU) 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. 2048. Pro algoritmizaci učení se pak může užít tzv. CUDA platforma, umožňující programovat v jazicích C++, resp. Fortran, která umožní díky velkému počtu dostupných výpočetních vláken vysoký stupeň paralelizace výpočtu, čímž zvýší výpočetní výkon strojového učení.

10. Strojové učení

Využití metod hlubokého učení

Rozpoznávání obrazu založené na hlubokém učení již překonalo schopnosti člověka. Poprvé se tak stalo v roce 2011 při rozpoznávání dopravního značení, poté v roce 2014 v rozpoznávání lidských tváří. Vozidla využívající hluboké učení dnes analyzují 360° kamerové záběry v reálném čase.

Pro rozpoznávání obrazu se dnes užívají **konvoluční neuronové sítě (CNN)**, které jsou specializovaným typem umělých neuronových sítí, které alespoň v jedné ze svých vrstev používají matematickou operaci zvanou konvoluce namísto obvyklého násobení matic. Jsou speciálně navrženy pro zpracování pixelových dat a používají se při rozpoznávání a zpracování obrazu.

V porovnání s jinými algoritmy klasifikace obrazu používají CNN relativně málo předzpracování. To znamená, že síť se učí optimalizovat filtry (nebo jádra) prostřednictvím **automatického učení**, zatímco v tradičních algoritmech jsou tyto filtry vytvářeny ručně. Tato nezávislost na předchozích znalostech a lidských zásazích do extrakce příznaků je velkou výhodou.

10. Strojové učení

Pro klasifikaci ručně psaných číslic se běžně využívá datový soubor MNIST. Jde o vzorník ručně psaných číslic obsahující 60 tisíc trénovacích položek a 10 tisíc testovacích položek. Relativně malý datový objem umožňuje testovat různé konfigurace systému pro rozpoznávání.

Automatické rozpoznávání řeči je nejočividnější aplikací metod hlubokého učení s širokým využitím. Rekurentní neuronové sítě jsou schopné zpracovávat i "velmi hluboké úlohy", které obsahují intervaly několika sekund trvající mluvené řeči oddělené několika tisíci diskrétních časových kroků, kde jeden krok odpovídá zhruba 10 ms.

Počátečního úspěchu v rozpoznávání řeči bylo dosaženo na úlohách malého rozsahu založených na souboru TIMIT (vzorník mluvené americké angličtiny od různých mluvčích doplněný transkripcí). Tato data obsahují po deseti větách od 630 mluvčích hovořících v osmi hlavních dialektech americké angličtiny.

Všechny dnes komerčně široce využívané systémy pro rozpoznávání řeči (Microsoft Cortana, Amazon Alexa, Google Now, Apple Siri a další) jsou založeny na hlubokém učení.

10. Strojové učení

Zpracování přirozeného jazyka

Klíčovou technikou v tomto oboru je vnoření slov (angl. *word embedding*). Tento přístup reprezentuje každé jednotlivé slovo jako pozici ve vícerozměrném prostoru relativně vůči dalším slovům obsaženým v datovém souboru. Využití vnoření slov na vstupní vrstvě umožňuje rekurentním neuronovým sítím provádět rozbor vět a frází efektivní vektorovou gramatikou. Rekurzivní autoenkodéry umožňují posoudit příbuznost vět a detekovat parafráze.

Nedávné pokroky zobecnily vnoření slov na vnoření celých vět (angl. *sentence embedding*).

Překladač Google a *Google Neural Machine Translation* využívají techniky hlubokého učení. Překladač Google díky tomu překládá významy vět namísto překladu jednotlivých frází. Podporuje přes sto jazyků a využívá angličtinu jako prostředníka pro překlad mezi libovolnými dvěma jazyky.

10. Strojové učení

Analýza lékařských snímků

Hluboké učení prokázalo konkurenceschopnost v medicínských aplikacích, jako je klasifikace rakovinných buněk, detekce lézí, vylepšení obrazu a dalších. Moderní nástroje využívající hluboké učení prokazují vysokou přesnost v detekci různých chorob a jejich využití dovoluje odborníkům zvyšovat efektivitu diagnóz.

Rekonstrukce obrazu

Jde o rekonstrukci obrazu na základě podkladových dat, která se k obrazu vážou, jako je například spektroskopické zobrazování nebo ultrazvukové zobrazování. Několik prací na toto téma demonstrovalo lepší výkon systémů využívajících hluboké učení ve srovnání s uvedenými metodami.

Detekce finančních podvodů

Hluboké učení se úspěšně uplatňuje v detekci finančních podvodů, detekci daňových úniků a praní špinavých peněz.

10. Strojové učení

Zpracování uměleckých děl

Úzce spjaté s obecným rozpoznáváním obrazu je využití hlubokého učení pro různé typy zadání týkající se uměleckých děl. Hluboké neuronové sítě se prokázaly schopnými např.:

- identifikovat období vzniku malby
- zachytit výtvarný styl daného díla a poté ho aplikovat na libovolnou fotografii či video
- generovat obrazy založené na náhodných vstupech (např. software DeepDream společnosti Google)

10. Strojové učení
