

Dopad umelých neurónových sietí na vyhľadávanie informácií*

Tomáš Meravý Murárik

Slovenská technická univerzita v Bratislave
Fakulta informatiky a informačných technológií
`xmeravymurarik@stuba.sk`

5 október 2023

Abstrakt

Tento článok sa zaoberá umelými neurónovými sieťami (Artificial Neural Networks - ANN), algoritmami inšpirovanými biologickými neurónovými sieťami, s cieľom priblížiť sa čo najbližšie k spracovaniu informácií v ľudskom mozgu. Skúma základnú architektúru ANN, ktorá pozostáva z troch vrstiev: vstupnej, skrytej a výstupnej vrstvy. Vstupná vrstva prijíma rôzne formy vstupných dát, ako sú obrazové pixely alebo textové hodnoty, zatiaľ čo skrytá vrstva funguje ako filter a transformuje tieto dáta do relevantných vzorov.

Článok tiež analyzuje význam aktivačných funkcií a váh v sieťach a diskutuje výhody a nevýhody použitia umelých neurónových sietí. Okrem toho sa zameriava na praktické aplikácie ANN v rôznych odvetviach. Tento článok je nápomocným prehľadom pre tých, ktorí chcú pochopiť fungovanie a využitie umelých neurónových sietí v širokom spektre aplikácií.

1 Úvod

[7]

Umelé neurónové siete (ANN) sú výpočtové systémy inšpirované skutočnými ľudskými nervovými sieťami. Tieto siete presahujú rámec obyčajných algoritmov a predstavujú významný míľnik v oblasti umelej inteligencie, kde stroje simulujú kognitívne procesy s pozoruhodnou presnosťou. V digitálnom veku sú ANN dôkazom našej schopnosti napodobniť a využiť neuveriteľný výpočtový výkon mozgu.

V posledných rokoch sa ANN stali nenahraditeľnou silou v stále rozvíjajúcej oblasti umelej inteligencie, najmä v oblasti hlbokého učenia. Ich vplyv sa prejavuje v rôznych oblastiach, od rozpoznávania obrazu a reči až po autonómne vozidlá a zdravotníctvo. ANN pripravili pôdu pre budúcnosť, v ktorej stroje dokážu vnímať, učiť sa a rozhodovať, čím pripravili pôdu pre prevratné aplikácie.

Tento článok skúma svet umelých neurónových sietí, odhaľuje ich vnútorné fungovanie, aplikácie a potenciálne výzvy. Pripravte sa na cestu neurónovými krajinami, ktoré poháňajú moderné technológie.

*Semestrálny projekt v predmete Metódy inžinierskej práce, ak. rok 2023/2024, vedenie:

2 Definícia a prehľad:

[3]

Architektúra ANN pozostáva z troch základných vrstiev: vstupnej vrstvy, skrytej vrstvy (vrstiev) a výstupnej vrstvy. Vstupná vrstva prijíma údaje, zatiaľ čo skryté vrstvy sa zaoberajú extrakciou funkcií a analýzou údajov. Výstupná vrstva poskytuje konečné výsledky. Toto usporiadanie, často označované ako viacvrstvový perceptrón (MLP), odráža schopnosť mozgu destilovať podstatné informácie z komplexných vstupov.

Jadrom ANN je aktivačná funkcia, kľúčový komponent, ktorý do siete vnáša nelinearitu. Umožňuje ANN identifikovať zložité vzťahy a transformovať surové údaje na použiteľné poznatky. ANN sa stávajú bystrými žiakmi tým, že upravujú silu spojenia, známu ako váhy, keď sa trénujú na údajoch. Za týmto procesom učenia stojí algoritmus spätného šírenia, ktorý umožňuje ANN minimalizovať chyby a časom zlepšovať svoje predpovede.

3 Typy neurónových sietí

[6] Umelé neurónové siete majú rôzne podoby, z ktorých každá je prispôbená konkrétnym úlohám a výzvam. Na tomto mieste rozoberieme niektoré z najvýznamnejších typov:

- Napájacie neurónové siete (FNN):
 - Sú to základné stavebné prvky ANN, pričom informácie prúdia jedným smerom, zo vstupnej vrstvy do výstupnej vrstvy. Používajú sa v mnohých aplikáciách vrátane rozpoznávania vzorov, klasifikácie a regresie.
- Konvolučné neurónové siete (CNN):
 - CNN sú vhodné na úlohy analýzy obrazu. Využívajú konvolučné vrstvy na automatické získavanie hierarchických vlastností z obrázkov, vďaka čomu sú nezameniteľné v úlohách, ako je rozpoznávanie obrazu a detekcia objektov.
- Rekurentné neurónové siete (RNN):
 - RNN majú pamäťové schopnosti, ktoré im umožňujú spracovávať sekvencie údajov. Vďaka tomu sú cenné pri spracovaní prirodzeného jazyka, rozpoznávaní reči a predpovedaní časových radov.
- Siete GRU (Gated Recurrent Unit):
 - GRU sú podobne ako LSTM navrhnuté jednoduchšie, ale stále účinné na modelovanie sekvencií. Sú obľúbené v aplikáciách, ako je rozpoznávanie reči a odporúčacie systémy.
- Generatívne adverzné siete (GAN):
 - GAN pozostávajú z dvoch neurónových sietí, generátora a diskriminátora, ktoré sú uzamknuté v súťaži. Táto architektúra je nápomocná pri generovaní realistických údajov, čo má uplatnenie pri vytváraní obrázkov, rozširovaní údajov a vytváraní deepfake.

- Samoorganizujúce sa mapy (SOM):
 - SOM sú formou nekontrolovaného učenia, ktoré je užitočné najmä na zmenšenie počtu dimenzií a zhlukovanie. Pomáhajú pri vizualizácii vysokorozmerných údajov a mapovaní príznakov.

Každý typ neurónovej siete je prispôsobený špecifickým údajom a problémovým oblastiam a pochopenie ich silných a slabých stránok je rozhodujúce pre výber najvhodnejšej architektúry.

4 Ako sa ANN učia

[1]

Pochopenie toho, ako sa ANN učia, je základom pre pochopenie ich schopností. Proces učenia zahŕňa rôzne aspekty:

- Stratové funkcie: Na vyčíslenie rozdielu medzi predpoveďami siete a skutočnými cieľovými hodnotami sa používajú stratové funkcie. Tieto funkcie vedú sieť k minimalizácii chýb počas tréningu.
- Algoritmus spätného šírenia: Spätné šírenie je základným kameňom tréningu ANN. Zahŕňa prenášanie informácií o chybách smerom dozadu cez sieť, čo umožňuje úpravy váh, ktoré minimalizujú chyby predpovedí. Je to kľúčová zložka na jemné doladenie výkonu siete.
- Optimalizačné techniky: Na aktualizáciu parametrov siete sa používajú optimalizačné algoritmy, ako napríklad algoritmy na určenie smeru a veľkosti aktualizácií váh, čím zabezpečujú konvergenciu siete k optimálnemu riešeniu.
- Regularizačné metódy: Nadmerné prispôbenie je v neurónových sieťach častým problémom. Na zabránenie prispôbenia sa siete šumu v údajoch a zlepšenie generalizácie sa používajú regularizačné techniky vrátane dropout a weight decay.

V tomto procese nemožno preceňovať význam údajov a tréningovej množiny údajov. Kvalita a množstvo údajov zohrávajú významnú úlohu pri výkonnosti siete. Tréningová množina údajov tvorí základ, na ktorom ANN získavajú svoje učiace a predikčné schopnosti.

5 Hlboké učenie a hlboké neurónové siete

[2] Hlboké učenie(deep learning), podoblasť strojového učenia(machine learning), sa vo veľkej miere opiera o hlboké neurónové siete. Pojem "hlboký" označuje prítomnosť viacerých skrytých vrstiev v sieti. Hlboké siete sú nevyhnutné na riešenie zložitých úloh, ktoré si vyžadujú vysokú úroveň abstrakcie. Vynikajú v extrakcii príznakov a v učení reprezentácie.

Prelomové objavy v oblasti hlbokého učenia posunuli túto oblasť do centra pozornosti. Hlboké neurónové siete dosiahli pozoruhodné výsledky v oblasti rozpoznávania obrazu a reči, porozumenia prirodzenému jazyku a dokonca aj pri hraní zložitých hier. Tieto pokroky vyvolali záujem a investície do vývoja modelov hlbokého učenia pre širokú škálu aplikácií.

6 Využitie ANN

[4] Umelé neurónové siete si našli cestu do množstva reálnych aplikácií v rôznych oblastiach:

- Rozpoznávanie obrazu a reči: ANN, vrátane RNN, sú základom systémov na rozpoznávanie reči: CNN spôsobili revolúciu v rozpoznávaní a analýze obrazu. Aplikácie siahajú od rozpoznávania tváre až po hlasových asistentov.
- Spracovanie prirodzeného jazyka: ANN poháňajú strojový preklad, chatboty a analýzu nálad. Umožňujú strojom porozumieť ľudskému jazyku a vytvárať ho, čím podporujú pokrok v úlohách súvisiacich s jazykom.
- Autonómne vozidlá: Hlboké učenie a ANN sú kľúčové pre autonómne vozidlá, ktorým pomáhajú vnímať a navigovať prostredie prostredníctvom rozpoznávania obrazu, analýzy údajov zo senzorov a rozhodovacích algoritmov.
- Zdravotníctvo: ANN prispievajú k analýze lekárskeho obrazov, objavovaniu liekov a predpovedaniu chorôb. Ponúkajú pohľad na údaje o pacientoch, čím pomáhajú pričasnej diagnostike a plánovaní liečby.
- Financie: Vo finančnom sektore sa ANN používajú na odhaľovanie podvodov, predpovedanie vývoja na burze a algoritmické obchodovanie. Obchodníkom a finančným inštitúciám poskytujú nástroje na rozhodovanie založené na údajoch.

Tieto aplikácie majú hlboký vplyv, zvyšujú produktivitu, presnosť a efektívnosť vo viacerých odvetviach. Keďže ANN naďalej napredujú, ich potenciálne prínosy sa môžu ešte viac rozšíriť.

7 Výzvy a obmedzenia

[5] Napriek svojim pozoruhodným schopnostiam majú ANN niekoľko obmedzení:

Nadmerné prispôsobenie: ako už bolo, vspomenuté v časti o Regularizačné metódach, ANN môžu byť náchylné na nadmerné prispôsobenie, keď dosahujú výnimočne dobré výsledky na trénovaných údajoch, ale nedokážu ich účinne zovšeobecniť na nepozorované údaje. Na zmiernenie nadmerného prispôsobenia sa používajú regularizačné techniky, ako napríklad dropout.

Požiadavky na údaje: ANN často vyžadujú veľké množstvo údajov na efektívne tréningy. Získanie dostatočného množstva označených údajov môže byť náročné, najmä v špecializovaných oblastiach alebo v prípade, že niektoré udalosti sú veľmi zriedkavé.

Interpretovateľnosť: Jedným z hlavných problémov ANN je ich nedostatočná interpretovateľnosť. Často sa považujú za "čierne skrinky", pretože môže byť náročné pochopiť, ako sa dospelo k určitým rozhodnutiam alebo predpovediam. To je významný problém, najmä v kritických aplikáciách, ako je zdravotníctvo a financie.

8 Záver

Na záver možno konštatovať, že umelé neurónové siete predstavujú základnú technológiu v súčasnej technike. Ich schopnosť učiť sa a modelovať zložité nelineárne vzťahy ich robí neoceniteľnými pre širokú škálu aplikácií. Keďže ANN riešia výzvy, ako je nadmerné prispôsobovanie a interpretovateľnosť, pripravujú pôdu pre väčšiu dôveru a prijatie.

Do budúcnosti sa ANN zosúlada s novými trendmi vrátane vysvetľujúcej umelej inteligencie, posilňovania učenia, transformátorov a kvantových neurónových sietí, aby sa ďalej rozšírili ich schopnosti a aplikácie. Tento vývoj sľubuje budúcnosť, v ktorej ANN budú naďalej formovať prostredie AI a strojového učenia a prinášať inovatívne riešenia a objavy. Ich význam v súčasnej technológii je nepopierateľný a cesta objavovania a napredovania v tejto oblasti pokračuje.

Literatúra

- [1] GeeksforGeeks. Artificial neural networks and its applications. <https://www.geeksforgeeks.org/artificial-neural-networks-and-its-applications/>, 2023.
- [2] Aurélien Géron. Neural networks and deep learning. <https://www.oreilly.com/library/view/neural-networks-and/9781492037354/ch01.html>, 2017.
- [3] Jahnvi Mahanta. Introduction to neural networks, advantages and applications. <https://www.kdnuggets.com/2017/07/introduction-neural-networks-advantages-applications.html>, 2017.
- [4] Amazon Web Services. What is a neural network? - artificial neural network explained - aws. <https://aws.amazon.com/what-is/neural-network/>. Accessed November 5, 2023.
- [5] Pratik Shukla. Introduction to artificial neural networks. <https://towardsdatascience.com/introduction-to-artificial-neural-networks-ac338f4154e5>, 2022.
- [6] Pratik Shukla and Roberto Iriondo. Main types of neural networks and its applications. <https://towardsai.net/p/machine-learning/main-types-of-neural-networks-and-its-applications-tutorial-734480d7ec8e>, 2020.
- [7] Guangyu Robert Yang and Xiao-Jing Wang. Artificial neural networks for neuroscientists: a primer. *Neuron*, 107(6):1048–1070, 2020.