Inteligencja w maszynach

Komputer zasługiwałby na miano inteligentnego, gdyby potrafił oszukać człowieka i wmówić mu, że jest człowiekiem. - Alan Turing

Inteligencja maszyn to ostatni wynalazek, jakiego ludzkość kiedykolwiek będzie musiała dokonać. - Nick Bostrom

Idea budowy inteligentnych maszyn istnieje od dawna. W starożytnym Egipcie i Chinach krążyły mity o robotach i ożywionych przedmiotach. Filozofowie rozważali hipotezę, że mechaniczni ludzie, sztuczne istoty i inne automaty istniały lub mogłyby istnieć w jakiejś formie. Uświadomienie sobie ludzkich zachowań poprzez imitację maszyn i uczynienie maszyn ludźmi posiadającymi ludzką inteligencję to długoterminowy cel ludzkości. Wraz z rozwojem komputerów cyfrowych inteligentne maszyny stały się coraz bardziej namacalne. Fala sztucznej inteligencji (AI) ogarnia świat. W tym rozdziale krótko opisano kluczowe wydarzenia w historii inteligentnych maszyn, omówiono szkoły techniczne, ważne algorytmy i przyszłe osiągnięcia.

Przed rokiem 1950

Różni teologowie, autorzy, matematycy, filozofowie i profesorowie rozważali techniki mechaniczne, maszyny liczące i systemy liczbowe, które mogą prowadzić do koncepcji mechanizacji istot nieludzkich i przekształcenia ich w ludzi. Na początku XVIII wieku, w swojej powieści „Podróże Guliwera", Jonathan Swift opisał urządzenie zwane „silnikiem", co stanowi jedno z pierwszych odniesień do współczesnego komputera ze sztuczną inteligencją. Dzięki temu urządzeniu, służącemu do doskonalenia wiedzy i operacji mechanicznych, nawet najmniej utalentowana osoba wydawała się być utalentowana. Pierwsze znane odniesienie do słowa „robot" znaleziono w sztuce science fiction „Rossum's Universal Robots" z 1921 roku, napisanej przez czeskiego dramatopisarza Karla Čapka. W sztuce tej występowały sztuczne, produkowane fabrycznie istoty zwane robotami. Od tego momentu ludzie zaczęli używać koncepcji „robota" i wdrażać ją w swoich badaniach, badaniach i rozwoju. Pierwszy film przedstawiający robota pojawił się w filmie science fiction z 1927 roku, wyreżyserowanym przez Fritza Langa. W filmie tym dziewczyna-robot zaatakowała miasto, siejąc spustoszenie w futurystycznym Berlinie. Film ten stał się inspiracją dla innych znanych postaci pozaludzkich, takich jak C-3PO w Gwiezdnych Wojnach. Pierwszym robotem zbudowanym w Japonii był Gakutensoku (w tłumaczeniu na język angielski „nauka o prawach natury"), stworzony przez japońskiego biologa Makoto Nishimurę w 1929 roku. Robot ten mógł poruszać głową i rękami oraz zmieniać mimikę twarzy, jak pokazano na rysunku.

W 1939 roku fizyk John Vincent Atanasoff i jego student Clifford Berry zbudowali komputer Atanasoffa-Berry'ego (ABC) na Uniwersytecie Stanowym Iowa. Komputer ABC mógł rozwiązywać do 29 równań liniowych jednocześnie i ważył ponad 317 kg. W książce informatyka Edmunda Berkeleya pt. „Giant Brains: Or Machines That Think", opublikowanej w 1949 r., zauważono, że wraz ze wzrostem możliwości przetwarzania dużych ilości informacji maszyny potrafią myśleć.

1940–1970: Narodziny sztucznej inteligencji (AI)

Rozwój technologii związanych ze sztuczną inteligencją (AI)

Rozwój technologiczny w latach 1940–1960 miał na celu połączenie funkcji zwierząt i maszyn. Norbert Wiener był pionierem cybernetyki, dążąc do ujednolicenia teorii sterowania i komunikacji zarówno u zwierząt, jak i u maszyn. Warren McCulloch i Walter Pitts opracowali matematyczny i komputerowy model neuronu biologicznego w 1943 roku . Wiele postępów w dziedzinie sztucznej inteligencji przyniosło owoce w latach 50. XX wieku. Claude Shannon, „ojciec teorii informacji", opublikował w 1950 roku artykuł zatytułowany „Programowanie komputera do gry w szachy", w którym opisał rozwój programu komputerowego grającego w szachy. W tym samym roku Alan Turing opublikował książkę „Maszyny obliczeniowe i inteligencja", w której zaproponował ideę gry w imitację, stawiając pytanie „Czy maszyny potrafią myśleć". Turing rozważał możliwość stworzenia myślących maszyn, które mogłyby prowadzić konwersację nieodróżnialną od rozmowy z człowiekiem. Propozycja ta stała się później Testem Turinga, mierzącym inteligencję maszyn. Test Turinga był pierwszą poważną propozycją i stał się ważnym elementem filozofii sztucznej inteligencji. Program komputerowy do gry w warcaby został opracowany w 1952 roku przez informatyka Arthura Samuela. Ten program jako pierwszy samodzielnie nauczył się grać w grę.

Propozycja koncepcji sztucznej inteligencji

W sierpniu 1956 roku w Dartmouth College w małym miasteczku Hannos naukowcy tacy jak New Hampshire, John McCarthy, Marvin Minsky (specjalista w dziedzinie sztucznej inteligencji i poznania), Claude Shannon, Allen Newell (informatyk) i Herbert Simon (laureat Nagrody Nobla w dziedzinie ekonomii) spotkali się, aby omówić temat, który wówczas wydawał się nieosiągalny: wykorzystanie maszyn do naśladowania ludzkiego uczenia się i innych aspektów inteligencji. Konferencja w Dartmouth trwała dwa miesiące. Chociaż nie osiągnięto powszechnego konsensusu, istniała nazwa dla tematu konferencji: sztuczna inteligencja. Dlatego rok 1956 stał się również pierwszym rokiem sztucznej inteligencji. Sztuczną inteligencję definiuje się jako zdolność maszyn do myślenia i uczenia się w sposób podobny do człowieka. Od pierwszego roku istnienia sztucznej inteligencji, badania i rozwój nad sztuczną inteligencją mają ponad 60-letnią historię. W tym okresie naukowcy z różnych dyscyplin i środowisk naukowych tworzyli własne interpretacje sztucznej inteligencji i przedstawiali odmienne punkty widzenia, co zaowocowało powstaniem różnych szkół akademickich. W tym okresie istniały trzy główne szkoły: symbolizm, koneksjonizm i behawioryzm, które wywarły ogromny wpływ na badania nad sztuczną inteligencją. Główną różnicą między tymi trzema szkołami jest opis różnych aspektów ludzkiej inteligencji (myślenia, mózgu, zachowania). Już w momencie pojawienia się koncepcji sztucznej inteligencji, rozpoczęła się walka kilku głównych frakcji zajmujących się sztuczną inteligencją. Zgodnie z metodologią symbolistów, sztuczna inteligencja powinna naśladować ludzką logikę w celu zdobywania wiedzy; koneksjoniści uważają, że duże zbiory danych i uczenie się są bardzo ważne; behawioryści uważają, że konkretne cele powinny być osiągane poprzez interakcję z otoczeniem.

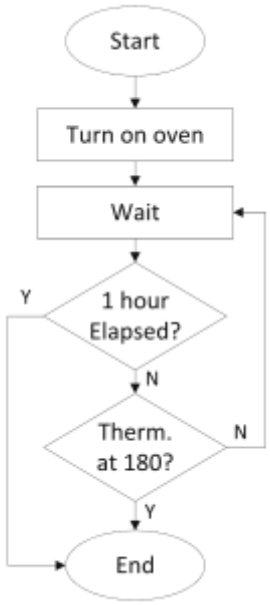
• Myśl (symbolizm): Wyraz świadomości naszego umysłu. Pochodzenie ludzkiej myśli, logiki abstrakcyjnej i emocji.

• Mózg (koneksjonizm): Niesamowita sieć neuronowa mózgu, która umożliwia myślenie.

• Zachowanie (behawioryzm): „odczuwająca-działająca” interakcja człowieka ze środowiskiem

Symboliczna sztuczna inteligencja

Po konferencji Dartmouth College w 1956 roku, uważanej za twórcę dyscypliny sztucznej inteligencji, lata 1956–1974 były złotym wiekiem sztucznej inteligencji. Pierwszym punktem kulminacyjnym sztucznej inteligencji jest symbolizm (znany również jako logicyzm, szkoła psychologii lub szkoła komputerowa). W pierwszych dekadach walki frakcyjnej frakcja symbolistów utrzymywała się w centrum uwagi, wyprzedzając rywali. Uczenie maszynowe, które dąży do koneksjonizmu, od dawna było pogardzane przez symbolistów w jego początkach. Od lat 50. do 70. XX wieku ludzie początkowo mieli nadzieję na osiągnięcie inteligencji maszynowej poprzez poprawę zdolności maszyn do logicznego rozumowania. Ogólnie rzecz biorąc, symbolizm zakłada, że podstawową jednostką ludzkiego myślenia są symbole, a szereg operacji opartych na symbolach stanowi proces poznania, dlatego ludzi i komputery można postrzegać jako systemy symboliczne posiadające zdolność logicznego rozumowania. Innymi słowy, komputery mogą symulować ludzką „inteligencję” poprzez różne operacje symboliczne. Rysunek 2 przedstawia przykładowy schemat blokowy w programie symbolicznym.



Ponieważ ludzkie procesy poznawcze i wyjaśnienie sztucznej inteligencji prezentowane przez tę szkołę są stosunkowo podobne i łatwo akceptowane przez wszystkich, symbolika od dawna dominuje w historii sztucznej inteligencji. Szkoła symbolistyczna uważa, że sztuczna inteligencja wywodzi się z logiki matematycznej, która rozwijała się dynamicznie od końca XIX wieku i była wykorzystywana do opisu inteligentnych zachowań w latach 30. XX wieku. Po pojawieniu się komputerów, system dedukcji logicznej został wdrożony na komputerach. Ludzie zawsze używali symboli do definiowania rzeczy (takich jak samochody), osób (takich jak nauczyciele), pojęć abstrakcyjnych (takich jak „miłość”), działań (takich jak bieganie) lub rzeczy, które nie istnieją fizycznie (takich jak mity). Jak omówiono wcześniej, uważa się, że zdolność do komunikacji za pomocą symboli czyni nas mądrzejszymi od innych zwierząt. Dlatego też dla wczesnych pionierów sztucznej inteligencji naturalne było założenie, że inteligencję można w zasadzie precyzyjnie opisać symbolami, a symboliczna sztuczna inteligencja zajęła centralne miejsce i stała się przedmiotem projektów badawczych nad sztuczną inteligencją. Ponadto wiele koncepcji i narzędzi w informatyce, jak na przykład programowanie obiektowe, jest efektem tych

wysiłków. Marvin Minsky, przedstawiciel symbolizmu, napisał książkę zatytułowaną „Perceptron”, która bezpośrednio zabiła sieci neuronowe i koneksjonizm, jak pokazano na rysunku 3.

Perceptron był siecią neuronową tamtej epoki. Minsky zaatakował w niej koneksjonizm. Twój perceptron nie potrafi nawet wykonać najprostszej operacji XOR. Jaki z niego pożytek [4]? W tym samym roku Minsky otrzymał również Nagrodę Turinga.

Osiągnięcia symbolicznej sztucznej inteligencji

John McCarthy opracował w 1958 roku Lisp, najpopularniejszy i nadal najpopularniejszy język programowania w badaniach nad sztuczną inteligencją . Termin „uczenie maszynowe” został ukuty przez Arthura Samuela, aby opisać zaprogramowanie komputera tak, aby grał w szachy lepiej niż osoba, która napisała program. Symbolika również daje pewne reprezentatywne rezultaty. Na przykład „teoretyk logiki”, którego pomysłodawcą był Allen Newell i inni, potrafi udowodnić 38 twierdzeń matematycznych zawartych w „Principia Mathematica” (później potrafi udowodnić wszystkie 52 twierdzenia). Niektóre rozwiązania są nawet bardziej pomysłowe niż te oferowane przez matematyków, na przykład koncepcje wyszukiwania heurystycznego. Innym przykładem jest architektura rozumowania ogólnego rozwiązywacza problemów (General Problem Solver) i koncepcje wyszukiwania heurystycznego zaproponowane przez Herberta Simona i innych, które mają dalekosiężny wpływ (na przykład AlphaGO czerpie z tej idei). Udanym przykładem symbolicznej sztucznej inteligencji (AI) były systemy eksperckie, które zostały zaprogramowane do symulowania osądu i zachowania człowieka lub organizacji posiadającej wiedzę ekspercką w danej dziedzinie [6]. „Maszyna wnioskowania” w tych systemach dostarcza odpowiedzi o wysokim poziomie wiedzy specjalistycznej, gdy zostanie o nie poproszona. Systemy eksperckie były szeroko stosowane w przemyśle. Znanym przykładem jest Deep Blue firmy IBM, który pokonał mistrza szachowego Kasparowa w 1997 roku. Japoński rząd hojnie finansował systemy eksperckie i inne przedsięwzięcia związane ze sztuczną inteligencją w ramach swojego projektu komputerowego piątej generacji (FGCP). Systemy eksperckie odegrały bardzo ważną rolę w promowaniu dobrobytu AI w XX wieku. Teoretycznie należą one również do wyników badań nad symboliką. Dzięki tym inspirującym historiom sukcesu sztuczna inteligencja zyskała bezprecedensowe zainteresowanie. Naukowcy są optymistycznie nastawieni do tego, że w pełni inteligentna maszyna powstanie w niecałe 20 lat. „Wnioskowanie jako przeszukiwanie” to popularny paradygmat, w którym pewne cele AI osiąga się poprzez przeszukiwanie labiryntu. Jednak po ponad dziesięciu latach badań powszechnie wiadomo, że chociaż zdolność logicznego rozumowania uległa poprawie, maszyna nie stała się mądrzejsza. Logika nie wydaje się być kluczem do otwarcia drzwi do inteligencji. Dlatego dodawana jest wiedza ludzka, czyli system

ekspercki, który do dziś rozwinął się w graf wiedzy. Główną trudnością związaną z tym paradygmatem jest to, że w przypadku wielu problemów liczba możliwych ścieżek rozwiązania dla sztucznej inteligencji jest astronomiczna. Już na tej podstawie niektóre problemy można rozwiązać, ale ich liczba jest nadal ograniczona.

Pierwsza zima AI

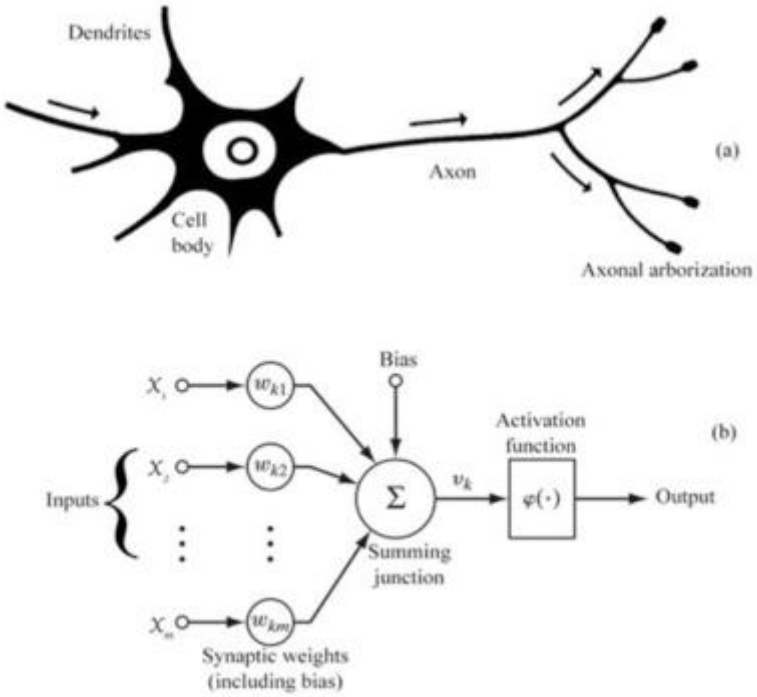
Lata 1974–1980 były pierwszą zimą AI. Ogromny optymizm wśród badaczy AI wzbudził wysokie oczekiwania. Gdy obiecane rezultaty się nie materializują, inwestycje i zainteresowanie AI znikają. Systemy ekspertowe najlepiej radzą sobie z problemami statycznymi, ale nie są naturalnym wyborem dla dynamicznych zagadnień w czasie rzeczywistym. Rozwój i utrzymanie stały się zatem niezwykle problematyczne. System ekspertowy może koncentrować się na wąskiej definicji inteligencji jako abstrakcyjnego rozumowania, daleko od zdolności modelowania złożoności świata. Inteligencja systemu ekspertowego ogranicza się do bardzo wąskiej dziedziny i trafniej byłoby powiedzieć, że jest on „żywym słownikiem". Główna trudność systemu ekspertowego polega na zdobywaniu i budowaniu wiedzy oraz realizacji mechanizmu wnioskowania. Dlatego naukowcy opracowali wiele teorii wokół tych trudności, takich jak rozumowanie łańcuchowe wsteczne i algorytm Rate. Graf wiedzy i eksploracja dużych zbiorów danych, z którymi zetknęliśmy się w ostatnich latach, są również mniej lub bardziej związane z rozwojem baz wiedzy. Awaria maszyny Lisp również ostudziła symbolikę. Lisp był wówczas językiem programowania powszechnie używanym w badaniach nad sztuczną inteligencją, a maszyna Lisp była komputerem zoptymalizowanym do uruchamiania programów w Lispie. W latach 80. XX wieku uczelnie studiujące sztuczną inteligencję kupowały takie maszyny, tylko po to, by odkryć, że nie potrafią tworzyć sztucznej inteligencji. Następnie pojawiły się komputery IBM PC i Mac, które były tańsze i wydajniejsze niż maszyny Lisp. Pod koniec lat 90. XX wieku, wraz z rozwojem nieudanego japońskiego inteligentnego komputera (piątej generacji) i upadkiem projektu Cyc Human Encyclopedia, kierowanego przez Uniwersytet Stanforda, sztuczna inteligencja ponownie wkroczyła w mroźną zimę. Termin AI stał się niemal tabu, a używa się łagodniejszych wariantów, takich jak „zaawansowane obliczenia". Ponadto dziedzina koneksjonizmu (lub sieci neuronowych) została niemal całkowicie zamknięta na 10 lat z powodu miażdżącej krytyki perceptronów przez Marvina Minsky'ego.

Koneksjonistyczna sztuczna inteligencja

Uczeni z nurtu koneksjonistycznego uważają, że sztuczna inteligencja wywodzi się z biomimikry, a zwłaszcza z badań nad modelami ludzkiego mózgu. Jej reprezentatywnym osiągnięciem jest model mózgu stworzony przez fizjologa McCullocha (McCulloch) i logika matematycznego Pittsa (Pitts) w 1943 roku, a mianowicie model MP. Opracowali oni nowe sposoby naśladowania struktury i funkcji ludzkiego mózgu za pomocą urządzeń elektronicznych. Wychodzą oni od neuronów, a następnie badają modele sieci neuronowych i modele mózgu, otwierając kolejną ścieżkę rozwoju sztucznej inteligencji.

Perceptron

Pierwszym przykładem sztucznej inteligencji inspirowanej koneksjonizmem był perceptron, wynaleziony przez psychologa F. Rosenblatta w latach 50. XX wieku. Inspiracją dla niego był sposób, w jaki neurony przetwarzają informacje w mózgu, jak pokazano na rysunku 4.



Neuron otrzymuje bodźce elektryczne lub chemiczne od innych neuronów. Jeśli suma wszystkich bodźców osiągnie określony próg, neuron się uaktywnia. Obliczając sumę swoich danych wejściowych, neuron przypisuje większą wagę danym wejściowym pochodzącym z silniejszych połączeń. Dostosowywanie siły połączeń między neuronami jest kluczowym elementem procesu uczenia się w mózgu. Analogicznie do neuronu, perceptron oblicza ważoną sumę swoich danych wejściowych i zwraca wartość 1, jeśli suma ta osiągnie określony próg. Jak określić wagi i próg w percepcji? W przeciwieństwie do symbolicznej sztucznej inteligencji, która ma wyraźne reguły ustalone przez programistę, percepcja uczy się tych wartości samodzielnie poprzez przykłady treningowe. Podczas treningu, jeśli wynik jest poprawny, zostanie nagrodzony; w przeciwnym razie zostanie ukarany. Jeśli perceptron zostanie rozszerzony poprzez dodanie warstw perceptronów, dzięki temu podejściu można rozwiązać szersze problemy. Ta nowa struktura, wielowarstwowa sieć neuronowa, stanowi podstawę dużej części współczesnej sztucznej inteligencji. Jednak w latach 50. i 60. XX wieku trenowanie sieci neuronowych było trudnym zadaniem, ponieważ nie istniały ogólne algorytmy do uczenia się wag i progów. Niestety, Frank Rosenblatt zginął tragicznie w wypadku na łodzi w 1971 roku w wieku 43 lat. Z powodu ograniczeń modeli teoretycznych, prototypów biologicznych i ówczesnych warunków technicznych, badania nad modelami mózgu pod koniec lat 70. i na początku lat 80. XX wieku podupadły. Bez wpływowych zwolenników i znacznego finansowania rządowego, badania nad sieciami neuronowymi i innymi systemami sztucznej inteligencji opartymi na koneksjonizmie praktycznie ustały. Zwłaszcza z powodu ostrej krytyki perceptronu przez Minsky'ego, frakcja koneksjonistów (lub zwolenników sieci neuronowych) od prawie dekady jest w kryzysie.

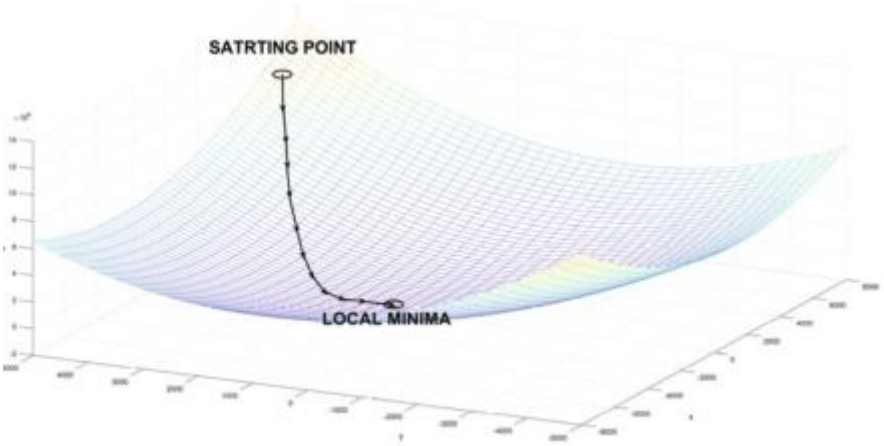
Uczenie maszynowe

Chociaż finansowanie koneksjonizmu gwałtownie spadło, niektórzy badacze koneksjonistyczni wytrwali w latach 70. i 80. XX wieku. Koneksjonizm odżył dopiero po opublikowaniu przez prof. Johna J. Hopfielda dwóch ważnych artykułów w 1982 i 1984 roku, w których zaproponował symulację sieci neuronowych w sprzęcie . W 1986 roku Rumelhart i inni zaproponowali algorytm propagacji wstecznej (BP) w sieciach wielowarstwowych. Od tego czasu koneksjonizm nabrał rozpędu, od modelu do algorytmu, od analizy teoretycznej do wdrożenia inżynierskiego, kładąc podwaliny pod wprowadzenie

sieci neuronowych na rynek. Uczenie maszynowe zyskało ogromną popularność od 2010 roku jako całkowite odejście od systemów eksperckich. Uczenie maszynowe nie wymaga reguł kodowania systemów eksperckich, lecz pozwala komputerom odkrywać je na podstawie ogromnych ilości danych. Uczenie maszynowe wpisuje się w koneksjonistyczne podejście do sztucznej inteligencji, które zasadniczo naśladuje działanie mózgu. W przeciwieństwie do symbolicznej sztucznej inteligencji (AI), która dąży do naśladowania koncepcji myślenia wyższego rzędu, koneksjonistyczna AI tworzy adaptacyjne sieci, które potrafią „uczyć się" i identyfikować wzorce na podstawie dużych ilości danych. Koneksjoniści uważają, że przy wystarczająco złożonych sieciach i wystarczającej ilości danych możliwe jest realizowanie bardziej zaawansowanych funkcji AI, równoważnych z rzeczywistym ludzkim umysłem.

Gradient spadkowy

Sieć neuronowa uczy się poprzez aktualizację swoich wag i progów. Standardowy algorytm uczenia się, który to umożliwia, nazywa się gradientem spadkowym. Koncepcja gradientu była już wielokrotnie omawiana . Przypomnijmy, że gradient jest jedynie miarą różnicy na dystansie (np. energii, masy, temperatury, informacji itp.). Ponieważ „natura nie znosi gradientu", a gradient oznacza niestabilność, system można ustabilizować poprzez zmniejszenie gradientu. Omawialiśmy ten proces w odniesieniu do zjawisk fizyki, chemii, biologii i człowieka. W tym procesie pojawia się inteligencja. Gradient w uczeniu maszynowym to różnica między rzeczywistym wynikiem działania maszyny a oczekiwanym wynikiem działania maszyny. Na przykład, jeśli chcesz zaprojektować inteligentną maszynę rozpoznającą kota, oczekiwanym wynikiem będzie „to jest kot", jeśli maszyna otrzyma zdjęcie kota. Jeśli rzeczywistym wynikiem maszyny jest „to jest pies", występuje gradient. Algorytm gradientu prostego służy do minimalizacji gradientu, tak aby rzeczywisty wynik był taki sam jak oczekiwany. Istnieje piękna analogia między algorytmem gradientu prostego a staczaniem się kamieni po zboczu doliny. Nawiasem mówiąc, to jeden z powodów, dla których uważam, że inteligencja pojawia się naturalnie w procesie stabilizacji wszechświata, tak naturalnie jak toczenie się kamieni. Różnicę między oczekiwanym a rzeczywistym wynikiem można modelować jako funkcję, która nazywa się funkcją kosztu (czasami określaną jako funkcja straty lub funkcja celu). Możemy myśleć o tej funkcji kosztu jako o pewnego rodzaju dolinie, a parametry (wagi i progi) sieci neuronowej określają położenie kamienia. Losowo wybieramy punkt startowy dla (wyimaginowanej) piłki, a następnie symulujemy ruch piłki podczas staczania się na dno doliny. Algorytm gradientu prostego działa poprzez obliczenie gradientu funkcji kosztu, z którego możliwe jest znalezienie kierunku „w dół" doliny, a następnie przesunięcie kamienia w dół (tj. zmiana parametrów sieci neuronowej) zbocza doliny, jak pokazano na rysunek 5.



Poprzez wielokrotne stosowanie tej reguły aktualizacji możemy „stoczyć się w dół wzgórza” i mieć nadzieję na znalezienie minimum funkcji kosztu. Innymi słowy, jest to reguła, której można użyć do uczenia się w sieci neuronowej, aż do osiągnięcia dna (tj. lokalnego minimum). Istnieje kilka wyzwań w bezpośrednim stosowaniu algorytmu gradientu prostego w praktyce. Jednym z nich jest niska prędkość, gdy liczba danych wejściowych do treningu jest bardzo duża. Aby przyspieszyć uczenie się, można zastosować stochastyczny gradient prosty. Pomysł polega na użyciu małej próbki losowo wybranych danych wejściowych do treningu, a nie wszystkich próbek.

Propagacja wsteczna

Kolejnym wyzwaniem związanym z metodą gradientu zstępującego jest efektywne obliczenie gradientu funkcji kosztu. Wyobraźmy sobie, że w naszej sieci znajduje się milion wag. Oznacza to, że aby obliczyć gradient, musimy obliczyć funkcję kosztu milion razy, co wymaga miliona przejść w przód przez sieć (dla każdego przykładu treningowego). Algorytm propagacji wstecznej unika powtarzających się podwyrażeń, a tym samym efektywnie oblicza gradient funkcji kosztu. W algorytmie propagacji wstecznej wagi sieci neuronowej są precyzyjnie dostrajane na podstawie współczynnika błędów uzyskanego z poprzedniego przebiegu. Prawidłowe zastosowanie tej metody może zmniejszyć współczynnik błędów. Po każdym przejściu w przód przez sieć algorytm wykonuje transfer wsteczny zgodnie z wagami i odchyleniami, dostosowuje parametry modelu i poprawia jego niezawodność. W szczególności błędy w wynikach sieci neuronowej są propagowane wstecz, aby przypisać odpowiednią winę do wag w sieci neuronowej. Stopniowo modyfikując wagi, błąd wyjściowy można zminimalizować do poziomu bliskiego zeru w miarę wykorzystywania coraz większej liczby próbek treningowych. Chociaż algorytm propagacji wstecznej pojawił się w latach 70. XX wieku, to dopiero David Rumelhart, Geoffrey Hinton i Ronald Williams w swojej słynnej pracy opublikowanej w 1986 roku opisali kilka sieci neuronowych, w których algorytmy działały szybciej niż wcześniejsze metody uczenia. Dlatego sieci neuronowe mogą być wykorzystywane do rozwiązywania problemów, które wcześniej nie były możliwe. Obecnie algorytm propagacji wstecznej jest podstawą uczenia się sieci neuronowych.

Uczenie nadzorowane

Zgodnie z metodą uczenia algorytmu, uczenie maszynowe można podzielić na trzy kategorie: uczenie nadzorowane, uczenie nienadzorowane i uczenie ze wzmocnieniem, jak pokazano w tabeli.

Porównanie różnych algorytmów uczenia maszynowego

Kryteria: Uczenie nadzorowane: Uczenie nienadzorowane: Uczenie przez wzmacnianie

Definicja: Uczenie się z wykorzystaniem danych oznaczonych ze wskazówkami: Uczenie się z wykorzystaniem danych nieoznaczonych bez wskazówek: Uczenie się poprzez interakcję ze środowiskiem

Typ danych: Dane oznaczone: Dane nieoznaczone: Brak danych predefiniowanych

Typ problemów: Klasyfikacja i regresja: Klastrowanie i asocjacja: Eksploatacja i eksploracja

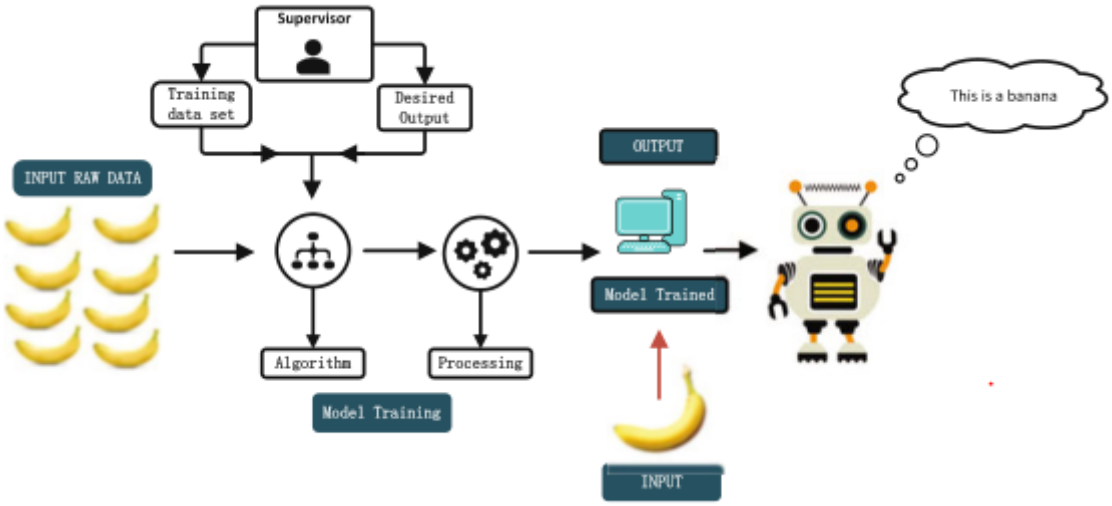
Nadzór: Z nadzorem: Bez nadzoru: Bez nadzoru

Algorytm: Regresja liniowa, regresja logistyczna, SVM, KNN itp.: K-Means, C-Means, Apriori itp.: Q-Learning, A3C itp.

Cel: Wyprowadzenie wyników: Odkrycie ukrytych wzorców: Optymalizacja długoterminowej nagrody

Zastosowanie: Rozpoznawanie obiektów, prognozowanie itp.: Rekomendacja, anomalia wykrywanie itp.: gry, samochody autonomiczne itp.

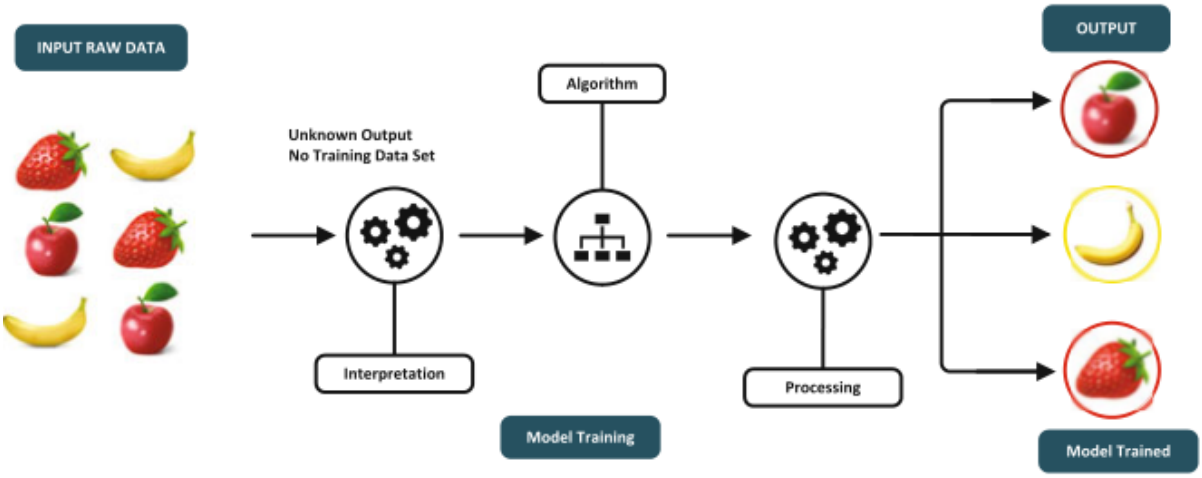
Uczenie nadzorowane odnosi się do uczenia się poprzez trenowanie modelu na oznaczonym zbiorze danych. Załóżmy, że jesteś uczniem siedzącym w klasie, a nauczyciel nadzoruje twoje zajęcia. Nauczyciel daje Ci zestaw pytań szkoleniowych. Po wykonaniu zestawu pytań szkoleniowych, nauczyciel informuje Cię, czy wykonałeś zadanie poprawnie. Uczenie nadzorowane ma podobną procedurę, w której zestaw danych z etykietami jest dostarczany z rozwiązaniem. Ułatwia to modelowi uczenie się. Rysunek 6 przedstawia przykład uczenia nadzorowanego.



Istnieją dwa rodzaje problemów, z którymi radzi sobie uczenie nadzorowane: problemy klasyfikacji i problemy regresji. W przypadku problemów klasyfikacji algorytm musi zaklasyfikować dane wejściowe (np. owoc) jako członka określonej grupy (np. jabłko, banan itp.). Problemy regresji są wykorzystywane do danych ciągłych, na przykład do przewidywania ceny akcji na giełdzie. Historia cen jest wysyłana do maszyny w celu szkolenia, a przyszła cena jest przewidywana przez algorytm.

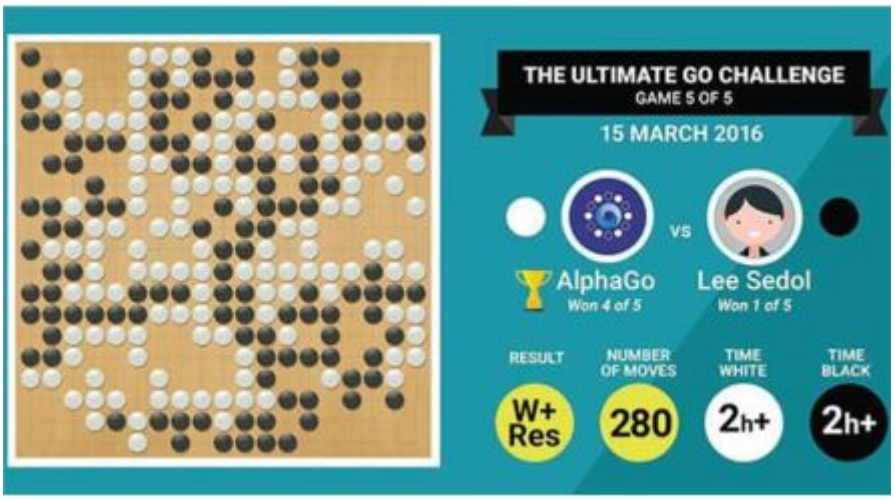
Uczenie się bez nadzoru

W przeciwieństwie do uczenia nadzorowanego, uczenie się bez nadzoru nie wymaga etykietowanych danych. Zamiast tego ma na celu znalezienie ukrytych relacji i wzorców w danych [16]. Uczenie się bez nadzoru to uczenie się samoorganizujące się. Maszyna otrzymuje dane i zostaje poproszona o wyszukanie ukrytych cech, a następnie musi je pogrupować w sposób, który ma sens. Typowym przykładem uczenia się bez nadzoru są algorytmy klastrowania, które pobierają zbiór danych i znajdują w nim grupy. Na przykład, powiedzmy, że chcemy posegmentować owoce na grupy, ale nie wiemy, jak najlepiej je zdefiniować. Algorytmy klastrowania mogą je zidentyfikować, jak pokazano na rysunku 7.



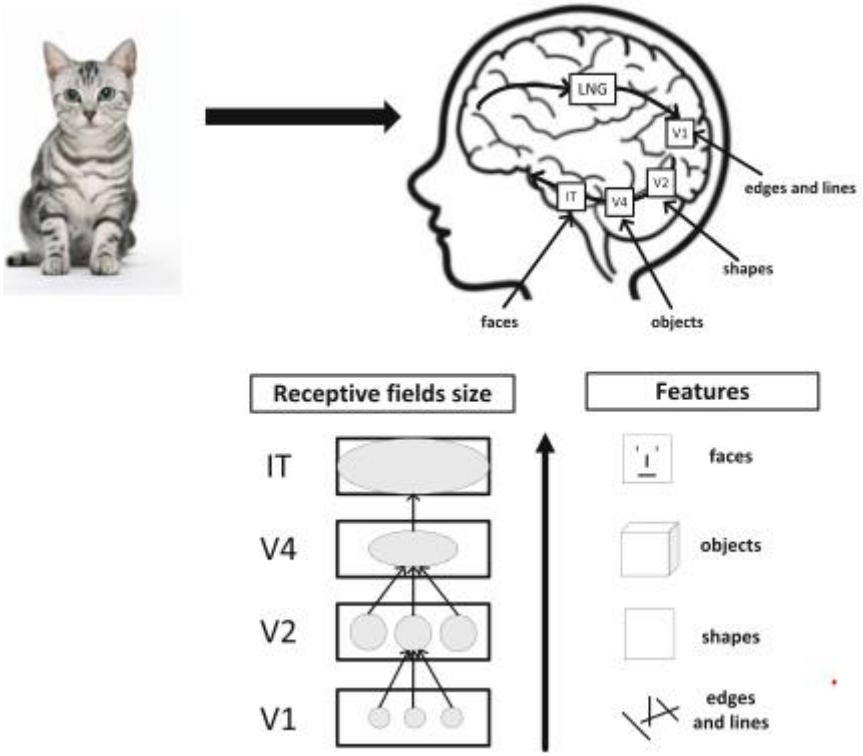
Głębokie uczenie

Spośród technik uczenia maszynowego, głębokie uczenie stało się najbardziej obiecującą techniką w wielu zastosowaniach, w tym w rozpoznawaniu głosu i obrazu. „Głębokie" w głębokim uczeniu odnosi się do głębokości warstw w sieci neuronowej. Sieć neuronowa składająca się z więcej niż trzech warstw, obejmująca dane wejściowe i wyjściowe, może być uznana za algorytm głębokiego uczenia. Sieci neuronowe stanowią podstawę algorytmów głębokiego uczenia. Chociaż badania nad głębokimi sieciami neuronowymi trwają od kilkudziesięciu lat, znaczące publiczne sukcesy głębokich sieci neuronowych w ostatnich latach wywołały nową falę zainteresowania sztuczną inteligencją. W 2011 roku Watson firmy IBM wygrał rozgrywki z dwoma mistrzami Jeopardy. W 2016 roku AlphaGO (sztuczna inteligencja Google specjalizująca się w grach Go) pokonała mistrza Europy (Fan Hui) i mistrza świata (Lee Sedol) ,



a następnie samą siebie (AlphaGo Zero). W 2020 roku AlphaFold rozwiązał jedno z największych wyzwań biologii: przewidzenie, w jaki sposób białka zwijają się z liniowego łańcucha aminokwasów w trójwymiarowe kształty, które umożliwiają im wykonywanie zadań życiowych. Niektóre z najskuteczniejszych głębokich sieci to te, których struktura naśladuje części mózgu, modelowane na podstawie odkryć w dziedzinie neuronauki. Od 1958 roku do końca lat siedemdziesiątych neurobiolodzy David H. Hubel i Torsten Wiesel pracowali razem nad badaniem właściwości pól

recepcyjnych neuronów w korze wzrokowej. Odkryli dwa główne typy komórek w pierwotnej korze wzrokowej. Pierwszy typ, komórki proste, reaguje na paski światła lub ciemności, gdy są umieszczone w określonych miejscach przestrzennych (tworząc krzywą dostrajania orientacji). Drugi typ, komórki złożone, ma mniej ścisłe profile odpowiedzi. Doszli do wniosku, że komórki złożone osiągają tę niezmienność poprzez łączenie sygnałów wejściowych z wielu komórek prostych [18]. Te dwie cechy (selektywność względem określonych cech i rosnąca niezmienność przestrzenna poprzez połączenia sprzężenia zwrotnego) stanowią podstawę sztucznych systemów wizyjnych. Ich praca kładzie podwaliny pod neuronaukę widzenia i dostarcza fundamentalnych spostrzeżeń na temat przetwarzania informacji w układzie wzrokowym. Za swoją pracę otrzymali w 1981 roku Nagrodę Nobla w dziedzinie fizjologii i medycyny. Rysunek 9 przedstawia ścieżkę bodźców wzrokowych i hierarchię wzrokową od oka do kory mózgowej.

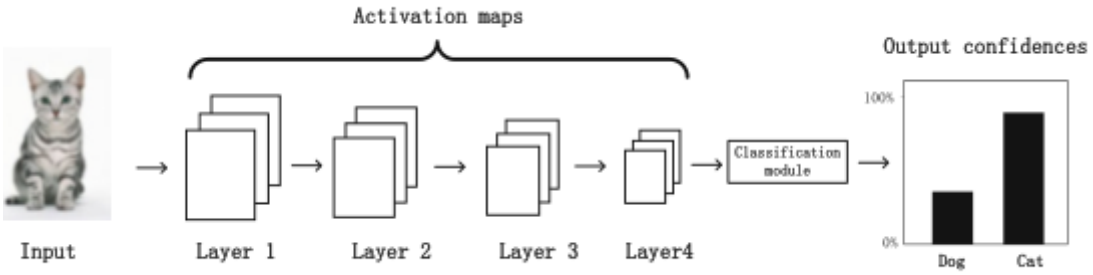


Wzdłuż ścieżki rozmiary pól recepcyjnych poszczególnych jednostek rosną wraz z postępem w warstwach sieci, tak jak w miarę przechodzenia od V1 do IT. Ponadto neurony w różnych warstwach tej hierarchii działają jak „detektory", które reagują na coraz bardziej złożone cechy sceny. Pierwsza warstwa wykrywa krawędzie i linie, następnie proste kształty złożone z tych krawędzi, aż po bardziej złożone kształty, obiekty i twarze. Zainspirowany odkryciami Hubela i Wiesela, japoński inżynier Kunihiko Fukushima opracował w latach 70. XX wieku jedną z pierwszych głębokich sieci neuronowych, zwaną „neokognicją", która po pewnym treningu z powodzeniem rozpoznawała odręcznie pisane cyfry. Chociaż neokognicja miała trudności z rozpoznawaniem złożonych treści wizualnych, stała się ona ważną inspiracją dla jednej z najpowszechniej stosowanych i najbardziej wpływowych głębokich sieci neuronowych – splotowych sieci neuronowych (CNN).

Sieci neuronowe splotowe

Sieci neuronowe splotowe zostały po raz pierwszy zaproponowane przez Yanna LeCuna w latach 80. XX wieku. Wytrenował on małą sieć neuronową splotową do rozpoznawania pisma odręcznego. Sieci

neuronowe splotowe poczyniły dalsze postępy w 1999 roku wraz z wprowadzeniem zbioru danych MNIST. Pomimo tych sukcesów, metody te stopniowo znikały ze środowiska badawczego, ponieważ uznawano je za trudne. Co więcej, wiele prac koncentrowało się na ręcznym projektowaniu cech wykrywanych na obrazach, w oparciu o przekonanie, że są one najbardziej informatywne. Po filtrowaniu na podstawie tych ręcznie stworzonych cech, uczenie się odbywa się dopiero w ostatnim etapie, czyli mapowaniu cech na klasy obiektów. Rysunek 10 przedstawia sieć neuronową splotową z 4 warstwami do rozpoznawania obrazów zwierząt.



Na rysunku 10 każda warstwa sieci neuronowej splotowej ma trzy nakładające się prostokąty. W rzeczywistej sieci neuronowej splotowej występuje wiele prostokątów. Te prostokąty reprezentują mapy aktywacji, podobne do układu wzrokowego mózgu odkrytego przez Hubera i Wiesela. Splotowe sieci neuronowe są trenowane od początku do końca za pomocą uczenia nadzorowanego, co umożliwia automatyczne generowanie cech w sposób najlepiej dopasowany do danego zadania.

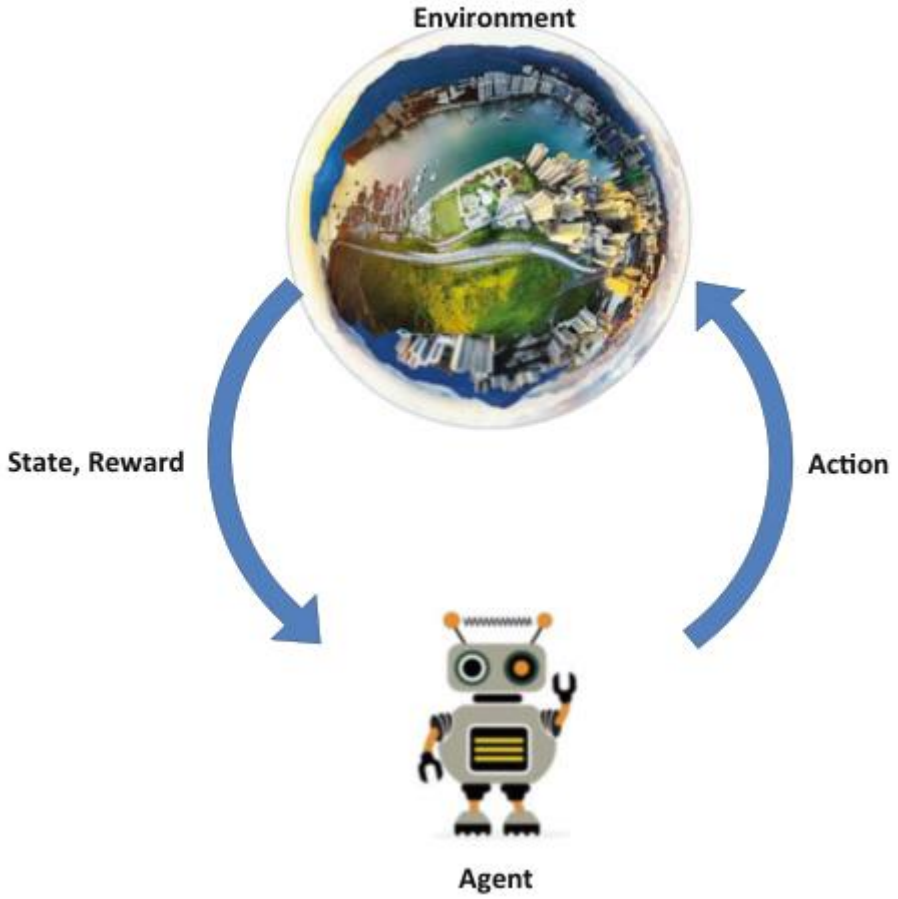
Behawioryzm

Inteligencja behawioralna

Behawioryzm to inteligentna metoda symulacji zachowań oparta na „percepcji-działaniu". Myśl cybernetyczna stała się ważną częścią ducha epoki już w latach 40. i 50. XX wieku, wpływając na wczesnych pracowników sztucznej inteligencji. Cybernetyka i systemy samoorganizujące się zaproponowane przez Weinera i McClocka i in. oraz cybernetyka inżynieryjna i cybernetyka biologiczna zaproponowane przez Qian Xuesena wywarły wpływ na wiele dziedzin. Cybernetyka łączy funkcjonowanie układu nerwowego z teorią informacji, teorią sterowania, logiką i komputerami. Wczesne prace badawcze koncentrowały się na symulacji inteligentnego zachowania i roli człowieka w procesie sterowania, na przykład badania nad systemami cybernetycznymi, takimi jak samooptymalizacja, samoadaptacja, samostabilizacja, samoorganizacja i samouczenie się. Zajmowano się również rozwojem „zwierząt cybernetycznych". W latach 60. i 70. XX wieku poczyniono pewne postępy w badaniach nad tymi systemami cybernetycznymi, zasiano ziarno inteligentnego sterowania i inteligentnych robotów, a w latach 80. narodziły się inteligentne systemy sterowania i inteligentnych robotów.

Uczenie się przez wzmocnienie

Uczenie się przez wzmocnienie (RL), znane również jako uczenie się przez wzmocnienie, uczenie się przez ocenę lub uczenie się przez wzmocnienie, jest jednym z paradygmatów i metodologii uczenia maszynowego. Problem strategii maksymalizacji zysków lub osiągnięcia określonych celów [20] przedstawiono na rysunku 11.



Uczenie się przez wzmocnienie jest inspirowane teorią behawiorystyczną w psychologii, czyli tym, w jaki sposób organizmy stopniowo kształtują antycypację bodźców pod wpływem stymulacji nagrodami lub karami ze środowiska i wytwarzają nawykowe zachowania, które mogą przynieść największe korzyści. Uczenie się przez wzmocnienie kładzie nacisk na to, jak działać w oparciu o środowisko, aby zmaksymalizować zamierzone korzyści. Dlatego też uczenie się przez wzmacnianie można zaliczyć do kategorii behawioryzmu. Uczenie się przez wzmacnianie można prześledzić do eksperymentów Pawłowa z warunkowaniem. W eksperymentach wielokrotne skojarzenie jednego bodźca z innym bodźcem bezwarunkowym, z nagrodą lub karą, może umożliwić osobnikom nauczenie się wywoływania reakcji warunkowej podobnej do reakcji bezwarunkowej, gdy bodziec jest prezentowany osobno. Najsłynniejszym przykładem warunkowania klasycznego jest warunkowanie śliną u psa Pawłowa. W tym eksperymencie, za każdym razem przed podaniem psu jedzenia, włączano czerwone światło i dzwoniono dzwonkiem. Po pewnym czasie, gdy tylko zadzwonił dzwonek lub zapaliło się czerwone światło, pies zaczynał się ślinić. Uczenie się przez wzmacnianie rozwijało się niezależnie w dwóch dziedzinach badań nad zachowaniem zwierząt i optymalnej kontroli, a ostatecznie zostało streszczone w procesie decyzyjnym Markowa (MDP) przez Richarda E. Bellmana . Proces decyzyjny Markowa to matematyczny model sekwencyjnego podejmowania decyzji, który służy do symulacji osiągalnych losowych strategii i nagród agenta w środowisku, w którym stan systemu ma właściwości Markowa. Proces decyzyjny Markowa został nazwany na cześć rosyjskiego matematyka Andrieja Markowa, na cześć jego badań nad łańcuchami Markowa. Uczenie ze wzmocnieniem (reinforcement learning) wykorzystuje metodę uczenia się podczas pobierania próbek środowiska i aktualizuje swój własny model po uzyskaniu próbek. Bieżący model jest wykorzystywany do kierowania działaniem agenta w jednym kroku, a następnie jest aktualizowany po nagrodzeniu kolejnego działania i powtarzany iteracyjnie, aż do osiągnięcia zbieżności. W tym procesie bardzo ważnym punktem jest „jak

wybrać kolejne działanie dla agenta, aby ulepszyć bieżący model, gdy bieżący model już istnieje", co obejmuje dwa bardzo ważne aspekty uczenia ze wzmocnieniem. Koncepcje: eksploatacja i eksploracja. Eksploatacja odnosi się do wybierania działań, które zostały wykonane w celu ulepszenia modelu znanych działań; eksploracja odnosi się do wybierania działań, które nie zostały wcześniej wykonane, aby zbadać więcej możliwości. Trzy najważniejsze cechy uczenia ze wzmocnieniem to:

1. Uczenie się zasadniczo odbywa się w formie zamkniętej pętli agenta i środowiska;

2. Agent nie wskazuje bezpośrednio, które działanie wybrać;

3. Seria działań i nagród może wpływać na proces uczenia się przez dłuższy czas.

Z uwagi na przełom w technologii głębokiego uczenia się w ostatnich latach, integracja uczenia wzmacniającego z głębokim uczeniem się umożliwiła dalsze wykorzystanie uczenia wzmacniającego. Na przykład, pozwól komputerowi nauczyć się grać w gry takie jak Go, StarCraft i tak dalej. Uczenie wzmacniające może również sprawić, że Twój program gry rozwinie się z całkowicie nieznajomego środowiska do mistrza w nim. AlphaGo z powodzeniem wykorzystał głębokie uczenie wzmacniające, aby pokonać profesjonalnych graczy w Go i stał się pierwszym robotem AI, który pokonał mistrza świata w Go [22]. AlphaGo został opracowany przez zespół DeepMind, firmy należącej do Google. Jego główną zasadą działania jest „głębokie uczenie wzmacniające". AlphaGo łączy podręczniki milionów ludzkich ekspertów Go z uczeniem wzmacniającym, aby się wyszkolić. W marcu 2016 roku AlphaGo rywalizował z mistrzem świata i profesjonalnym graczem z dziewięcioma danami, Lee Sedolem, w bitwie człowiek-maszyna w Go, wygrywając partię z wynikiem 4 do 1. Od końca 2016 do początku 2017 roku program korzystał z konta „Master" jako konta zarejestrowanego na chińskiej stronie szachowej, aby rywalizować z dziesiątkami mistrzów Go z Chin, Japonii i Korei Południowej w szybkich meczach szachowych, nie ponosząc ani jednej porażki w 60 kolejnych rundach. W maju 2017 roku, podczas szczytu Wuzhen Go w Chinach, AlphaGo zmierzył się z mistrzem świata numer 1 w Go, Ke Jie, wygrywając w sumie 3:0. W świecie Go uznaje się, że siła szachowa AlphaGo przewyższyła najwyższy poziom profesjonalnego Go. W światowym rankingu profesjonalnych graczy Go, opublikowanym przez stronę GoRatings, AlphaGo przewyższył Ke Jie, pierwszego gracza w rankingu. 27 maja 2017 roku, po bitwie człowiek-maszyna pomiędzy Ke Jie i AlphaGo, zespół AlphaGo ogłosił, że AlphaGo nie będzie już brał udziału w rozgrywkach Go. 18 października 2017 roku zespół DeepMind ogłosił najmocniejszą wersję AlphaGo o nazwie kodowej AlphaGo Zero. Na tej podstawie możliwości AlphaGoZero zostały jakościowo ulepszone. Główną różnicą jest to, że nie wymaga już danych od człowieka. Oznacza to, że nie ma on kontaktu z ludzkimi rekordami szachowymi. Zespół badawczo-rozwojowy po prostu pozwolił mu swobodnie grać w szachy na szachownicy, a następnie sam grał. AlphaGoZero przekształcił się w nauczyciela, wykorzystując nową metodę uczenia się przez wzmacnianie. System początkowo nie wiedział nawet, czym jest Go, ale zaczął od pojedynczej sieci neuronowej i grał przeciwko sobie, wykorzystując potężny algorytm wyszukiwania sieci neuronowej. Wraz ze wzrostem liczby graczy, sieć neuronowa stopniowo dostosowuje się, aby poprawić swoją zdolność przewidywania kolejnego kroku i ostatecznie wygrywa grę. Co więcej, wraz z pogłębianiem treningu, zespół AlphaGoZero odkrył, że AlphaGoZero samodzielnie odkrył również zasady gry. I opracował nowe strategie, wnosząc nowe spostrzeżenia do starożytnej gry Go. Metoda uczenia się przez wzmacnianie jest uniwersalna, dlatego była badana w wielu innych dziedzinach, takich jak cybernetyka, teoria gier, badania operacyjne, teoria informacji, metody optymalizacji symulacji, uczenie się systemów wieloagentowych, inteligencja roju, statystyka i algorytmy genetyczne.

Walka i unifikacja różnych szkół sztucznej inteligencji

Już w momencie zaproponowania koncepcji sztucznej inteligencji, rozpoczęła się walka wyżej wymienionych szkół. Symboliści uważają, że inteligentne maszyny powinny naśladować logiczny

sposób myślenia ludzi, aby zdobywać wiedzę. W oczach koneksjonistów duże zbiory danych i uczenie się poprzez trening są ważne. Behawioryści uważają, że sztuczna inteligencja powinna osiągać określone cele poprzez interakcję agentów z otoczeniem. W historii, zima sztucznej inteligencji jest mniej lub bardziej związana z walką kilku głównych szkół. Marzeniem ludzi zawsze było opisanie i zbadanie sztucznej inteligencji za pomocą jednolitej teorii. Aby stawić czoła problemom, które sztuczna inteligencja musi rozwiązać w świecie rzeczywistym, agenci muszą umieć radzić sobie ze złożonością i niepewnością. Symboliczna sztuczna inteligencja koncentruje się przede wszystkim na kwestiach złożoności, abstrahując złożony świat za pomocą logicznych relacji i wiedzy o świecie, podczas gdy koneksjonistyczna i behawioralna sztuczna inteligencja koncentruje się przede wszystkim na kwestiach niepewności, wykorzystując reprezentacje probabilistyczne. Jednak symboliczna sztuczna inteligencja opiera się na ograniczonej wiedzy ludzkiej i nie jest w stanie skutecznie odkryć subtelnej logiki i nieznanych praw. Symboliczna sztuczna inteligencja jest często zbyt krucha, aby poradzić sobie z niepewnością i szumem, które występują w wielu aplikacjach. Koneksjonistyczna i behawioralna sztuczna inteligencja często zmaga się ze złożonymi koncepcjami i relacjami. Gdy struktura sieci neuronowej jest zbyt prosta, istnieje ryzyko niedopasowania; gdy struktura sieci neuronowej jest zbyt złożona, występuje przedopasowanie. Trenowanie koneksjonistycznej i behawioralnej sztucznej inteligencji wymaga dużej ilości danych. Czarna skrzynka koneksjonistycznej i behawioralnej sztucznej inteligencji tworzy niewytłumaczalność, która sprawia, że systemy o znaczeniu krytycznym (takie jak autonomiczne pojazdy) nie mogą polegać na koneksjonistycznej i behawioralnej sztucznej inteligencji. Aby poradzić sobie ze złożonością i niepewnością występującą w większości rzeczywistych problemów, potrzebujemy połączenia symbolicznej, koneksjonistycznej i behawioralnej sztucznej inteligencji. Samodzielne rozwiązanie nie zapewni funkcjonalności niezbędnej do obsługi aplikacji AI. Obecnie implementacja symbolicznej sztucznej inteligencji jest nadal znacznie szersza niż koneksjonistycznej i behawioralnej sztucznej inteligencji, ponieważ wszystkie podstawowe funkcje współczesnych komputerów, wszystkie operacje matematyczne, całe tradycyjne oprogramowanie i aplikacje wykorzystują logikę symboliczną, nawet jeśli zaawansowane funkcje są sterowane statystycznie. W przyszłości te szkoły będą musiały osiągnąć pełną konwergencję, ponieważ większość zastosowań sztucznej inteligencji wymaga zarówno mocy ekspresji sztucznej inteligencji symbolicznej, jak i probabilistycznej solidności koneksjonistycznej i behawioralnej. Niestety, podział między symboliczną sztuczną inteligencją a koneksjonistyczną i behawioralną jest głęboki. Jak pokazano w tym rozdziale, sięga on początków tej dziedziny i jest nadal bardzo widoczny. Potrzebne są dalsze badania, aby ostatecznie połączyć te różne podejścia do sztucznej inteligencji.

Sztuczna inteligencja ogólna

Optymistyczny pogląd

Jestem pewien, że słyszeliście Państwo zaskakujące wieści o sztucznej inteligencji, która właśnie osiągnęła coś, co robią ludzie, a nawet więcej. Dzięki ostatnim postępom w dziedzinie sztucznej inteligencji, przepaść między inteligencją ludzką a sztuczną inteligencją wydaje się szybko zmniejszać. Wiadomości takie jak te i filmy science fiction pozwalają nam wierzyć, że rozwój ogólnej sztucznej inteligencji (AGI) lub sztucznej superinteligencji może być niedaleką przyszłością. AGI to hipotetyczna zdolność inteligentnego agenta do zrozumienia lub nauczenia się dowolnego zadania intelektualnego, które może wykonać człowiek. Wielu ekspertów jest optymistycznie nastawionych do AGI. Jedna z najsłynniejszych prognoz pochodzi od słynnego wynalazcy i futurologa Raya Kurzweila, który wpadł na pomysł osobliwości sztucznej inteligencji. W niedalekiej przyszłości, gdy komputery będą zdolne do samodoskonalenia i autonomicznego uczenia się, szybko osiągną i przewyższą ludzką inteligencję. Google zatrudniło go w 2012 roku, aby pomógł w realizacji tej wizji. Wszystkie przewidywania Kurzweila opierają się na idei „postępu wykładniczego” w wielu dziedzinach nauki i technologii,

zwłaszcza w informatyce. Na przykład, zgodnie z prawem Moore'a, liczba komponentów w chipie komputerowym podwaja się mniej więcej co 18 miesięcy, co skutkuje mniejszymi (i tańszymi) komponentami oraz wykładniczym wzrostem szybkości obliczeniowej i pamięci.

Pesymistyczny punkt widzenia

W rzeczywistości doniesienia o tym, że komputery właśnie zdołały coś zrobić lepiej niż ludzie, pojawiają się w historii komputerów. W latach 40. XX wieku komputery zastąpiły ludzi w obliczaniu trajektorii pędzącego pocisku i stały się nadludźmi. Było to pierwsze z wielu wąskich zadań, w których komputery osiągnęły sukces. W historii sztucznej inteligencji wielu praktyków było przesadnie optymistycznych. Na przykład, w 1965 roku pionier sztucznej inteligencji Herbert A. Simon stwierdził: „w ciągu dwudziestu lat maszyny będą zdolne do wykonywania każdej pracy, jaką może wykonać człowiek". Japoński komputer piątej generacji z 1980 roku miał dziesięcioletnią oś czasu, której celem było „prowadzenie luźnych rozmów". Jednakże, chociaż ostatnie postępy w dziedzinie sztucznej inteligencji (AI) podkreślają jej zdolność do wykonywania zadań z większą skutecznością niż ludzie, to generalnie nie jest ona inteligentna. W przypadku pojedynczej funkcji (np. gry Go) jest nadzwyczaj dobra, nie posiadając jednocześnie żadnych zdolności do robienia czegokolwiek innego. Dlatego, chociaż aplikacja AI może być równie skuteczna jak dorosły człowiek w wykonywaniu jednego konkretnego zadania, może przegrać z małym dzieckiem w konkurencji z innymi zadaniami. Na przykład, chociaż systemy wizyjne potrafią interpretować informacje wizualne, nie potrafią zastosować tej umiejętności do innych zadań. Z kolei człowiek, choć czasami mniej biegły w wykonywaniu określonej funkcji, może wykonywać szerszy zakres funkcji niż jakakolwiek z istniejących obecnie aplikacji AI. Niedawny sukces głębokiego uczenia się wynika nie tyle z nowych przełomów w dziedzinie AI, co z dostępności ogromnych ilości danych z Internetu i postępu w sprzęcie komputerowym, zwłaszcza w procesorach graficznych (GPU). Yann LeCun zauważył: „Rzadko się zdarza, żeby technologia istniejąca od 20, 25 lat – praktycznie niezmieniona – okazała się najlepsza. Szybkość, z jaką ludzie ją przyjęli, jest wręcz zdumiewająca. Nigdy wcześniej czegoś takiego nie widziałem". Dane treningowe mają znaczący wpływ na uczenie głębokie. W zasadzie, mając nieskończone dane, systemy uczenia głębokiego są wystarczająco wydajne, aby reprezentować dowolne skończone, deterministyczne „odwzorowanie" między dowolnym zestawem danych wejściowych a zestawem odpowiadających im danych wyjściowych. Jeden z najbardziej rozbudowanych modeli uczenia głębokiego, zaprojektowany do generowania języka zbliżonego do ludzkiego, znany jako GPT-3, czyli Generative Pre-trained Transformer trzeciej generacji, to model uczenia maszynowego oparty na sieci neuronowej, trenowany z wykorzystaniem danych internetowych do generowania dowolnego rodzaju tekstu. Sieć neuronowa uczenia głębokiego GPT-3 ma ponad 175 miliardów parametrów uczenia maszynowego, wymaga ilości energii równoważnej rocznemu zużyciu energii przez 126 duńskich domów i generuje ślad węglowy równoważny przejechaniu 700 000 kilometrów samochodem w ramach jednej sesji treningowej. Dla porównania, ludzki mózg pracuje z mocą 20 watów. To wystarcza, aby pokryć całą naszą zdolność myślenia. Sztuczna inteligencja potrzebuje niewiarygodnej ilości energii, aby rozpoznać zdjęcie kota spośród milionów obrazów. Aby rozwiązać ten problem, potrzebuje całych centrów danych, które muszą być chłodzone. Gdybyśmy chcieli użyć sztucznej inteligencji do odtworzenia wszystkiego, do czego zdolny jest ludzki mózg, potrzebowalibyśmy ogromnej liczby elektrowni jądrowych, aby zapewnić niezbędną energię, zakładając, że w ogóle byłoby to możliwe. Gary Marcus, badacz sztucznej inteligencji i profesor na Wydziale Psychologii Uniwersytetu Nowojorskiego, jest pesymistą co do sztucznej inteligencji ogólnej (AGI), ponieważ techniki głębokiego uczenia blokują sposoby przedstawiania związków przyczynowo-skutkowych (takich jak między chorobami a ich objawami) i prawdopodobnie napotkają trudności w pozyskiwaniu abstrakcyjnych idei. Nie mają oczywistych sposobów przeprowadzania logicznych wnioskowań i wciąż daleko im do integracji abstrakcyjnej wiedzy, takiej jak informacje o tym, czym są obiekty, do czego służą i jak są zazwyczaj używane. Uważa

on, że „ogólna sztuczna inteligencja na poziomie ludzkim nie poczyniła prawie żadnego postępu”. Jak daleko jesteśmy od stworzenia ogólnej sztucznej inteligencji? „Weź swoją ocenę, podwój ją, potrój ją, poczwórz. Wtedy” – to komentarz Orena Etzioniego, dyrektora Allen Institute for AI. Andrej Karpathy, starszy dyrektor ds. AI w Tesli, powiedział: „Jesteśmy naprawdę, naprawdę daleko”. Ten pesymistyczny pogląd podziela wielu innych badaczy, w tym Melanie Mitchell, autorka książki Artificial Intelligence: A Guide for Thinking Humans.

Natura inteligencji i nauka o inteligencji

Niezależnie od tego, czy chodzi o rozwiązanie problemu kilku głównych frakcji sztucznej inteligencji, czy o ogólną sztuczną inteligencję, ludzie z niecierpliwością oczekują na odkrycie istoty inteligencji. Musimy nie tylko budować systemy sztucznej inteligencji do widzenia i rozumienia języka naturalnego, ale także zrozumieć naturę inteligencji. Obecne prace nad sztuczną inteligencją koncentrują się głównie na projektowaniu nowych produktów, systemów i pomysłów. Są to głównie prace inżynieryjne, a brakuje zgłębiania natury inteligencji i nauki o inteligencji. Dlatego obecnie sztuczna inteligencja jest przede wszystkim technologią, a nie nauką. To, co muszą zrobić badacze sztucznej inteligencji, to budować i projektować potężne inteligentne systemy. Jeśli system działa dobrze, staramy się dowiedzieć, dlaczego działa dobrze – to jest nauka. Naukowiec musi wymyślić nowe koncepcje opisujące świat, a następnie wykorzystać metodę naukową do zbadania zasad wyjaśniających system, które są również dwoma aspektami sztucznej inteligencji. Badanie sztucznej inteligencji to nie tylko problem techniczny, ale także problem naukowy. Ostateczne pytanie brzmi, próbujemy dowiedzieć się, czym jest inteligencja. W przypadku silnika parowego nowe wynalazki napędzają badania teoretyczne. Ponad sto lat po tym, jak naukowcy wynaleźli silnik parowy, narodziła się termodynamika, a termodynamika jest zasadniczo podstawą wszystkich nauk ścisłych i przyrodniczych. Innym przykładem jest wynalezienie samolotu. Clément Ader, pionier francuskiego lotnictwa pod koniec XIX wieku, był genialnym inżynierem, który w latach 90. XIX wieku zbudował samoloty, które mogły faktycznie startować o własnej mocy. Zbudował samolot 30 lat przed braćmi Wright. Ale jego samolot miał kształt ptaka i brakowało mu sterowności. Tak więc po starcie samolot rozbił się po przelocie 15 metrów na wysokości około 15 centymetrów nad ziemią. Powodem jest to, że rozważał tylko bionikę, ale tak naprawdę nie rozumiał jej zasady. Rysunek 12 przedstawia samolot przypominający ptaka, Avion III, zaprojektowany przez Clémenta Adera, francuskiego pioniera lotnictwa z końca XIX wieku.

Samolot Adaira był pełen wyobraźni, a on sam był geniuszem w projektowaniu silników, ale z powodu braku teoretycznego wsparcia dla aerodynamiki jego projekt nie rozwinął się zbytnio. To ciekawa lekcja dla każdego, kto próbuje czerpać inspirację z biologii – musimy również zrozumieć, jakie jest uzasadnienie. W biologii istnieje wiele szczegółów, które są nieistotne. Bracia Wright w 1903 roku, a wcześniej Clément, wynaleźli samolot. Ponad trzy dekady później Theodore von Kármán odkrył teorię

aerodynamiki. W tym przykładzie wynalezienie samolotu było co najmniej tak samo ważne, jak aerodynamika. Na przykład w przypadku sztucznej inteligencji, głębokie uczenie się działa bardzo dobrze, jest to wynalazek, wkład, bardzo potężny system sztucznej inteligencji. Oczywiście musimy zbadać, dlaczego głębokie uczenie się działa tak dobrze, naturę inteligencji i naukę o inteligencji.