

Głęboka nauka. Rewolucja w AI

Weźmy dowolny stary problem z klasyfikacją, w którym masz dużo danych, a zostanie on rozwiązany przez głębokie uczenie. Będą tysiące zastosowań głębokiego uczenia się. - Geoffrey Hinton

Angielsko-kanadyjska psycholog kognitywna i informatyk Fei-Fei Li, która uzyskała tytuł licencjata z fizyki w Princeton w 1999 roku z wysokim wyróżnieniem oraz doktorat z elektrotechniki w Caltech w 2005 roku, skupiła się na tworzeniu modeli AI. Miała jednak poważne wyzwanie: znalezienie wysokiej jakości zbiorów danych. Początkowo przyglądała się tworzeniu ich ręcznie, na przykład ze studentami, którzy pobierali obrazy z Internetu. Ale proces był zbyt powolny i żmudny. Pewnego dnia student wspomniiał Li, że Mechanical Turk firmy Amazon.com, usługa online wykorzystująca crowdsourcing do rozwiązywania problemów, może być dobrym sposobem na skalowanie procesu. Pozwoliłoby to na szybkie i dokładne oznakowanie danych. Li spróbował i wyszło całkiem nieźle. Do 2010 roku stworzyła ImageNet, który zawierał 3,2 miliona obrazów w ponad 5200 kategoriach. Jednak spotkał się z chłodną reakcją społeczności akademickiej. Ale to nie powstrzymało Li. Kontynuowała niestrudzoną pracę nad ewangelizacją zbioru danych. W 2012 roku ogłosiła konkurs, aby zachęcić naukowców do tworzenia bardziej efektywnych modeli i przesuwania granic innowacji. Okazałby się przełomem, a konkurs stałby się corocznym wydarzeniem. W pierwszym konkursie profesorowie z University of Toronto-Geoffrey Hinton, Ilya Sutskever i Alex Krizhevsky wykorzystali zaawansowane algorytmy głębokiego uczenia. A wyniki były wyjątkowe. Zbudowany przez nich system, nazwany AlexNet, pokonał wszystkich pozostałych uczestników o 10,8%. To nie był przypadek. W następnych latach głębokie uczenie się nadal wykazywało przyspieszone postępy z ImageNet. W tej chwili wskaźnik błędów dla głębokiego uczenia się wynosi zaledwie 2%, co jest lepsze niż u ludzi. Nawiasem mówiąc, Li od tego czasu została profesorem w Stanford i współdyrektorem szkolnego laboratorium sztucznej inteligencji. Jest także głównym naukowcem Google zajmującym się sztuczną inteligencją i uczeniem maszynowym. Nie trzeba dodawać, że ilekroć ma teraz nowe pomysły, ludzie jej słuchają. Przyjrzymy się głębokiemu uczeniu, które jest zdecydowanie najgorętszym obszarem sztucznej inteligencji. Doprowadziło to do znacznych postępów w obszarach takich jak autonomiczne samochody i wirtualni asystenci, tacy jak Siri. Tak, głębokie uczenie się może być skomplikowanym tematem, a dziedzina ciągle się zmienia. Ale przyjrzymy się głównym koncepcjom i trendom – bez wchodzenia w szczegóły techniczne.

Różnica między głębokim uczeniem a uczeniem maszynowym

Często dochodzi do zamieszania między głębokim uczeniem a uczeniem maszynowym. I to jest rozsądne. Oba tematy są dość złożone i mają wiele podobieństw. Aby zrozumieć różnice, przyjrzymy się najpierw dwóm wysokopoziomowym aspektom uczenia maszynowego i ich powiązaniu z uczeniem głębokim. Przede wszystkim, chociaż oba zwykle wymagają dużej ilości danych, typy są generalnie różne. Weźmy następujący przykład: Załóżmy, że mamy zdjęcia tysięcy zwierząt i chcemy stworzyć algorytm do wyszukiwania koni. Cóż, uczenie maszynowe nie może analizować samych zdjęć; zamiast tego dane muszą być oznakowane. Algorytm uczenia maszynowego zostanie następnie przeszkolony do rozpoznawania koni w procesie znanym jako uczenie nadzorowane. Nawet jeśli uczenie maszynowe prawdopodobnie przyniesie dobre wyniki, nadal będą miały ograniczenia. Czy nie byłoby lepiej przyjrzeć się pikselom samych obrazów i znaleźć wzory? Zdecydowanie. Ale aby to zrobić z uczeniem maszynowym, musisz użyć procesu zwanego wyodrębnianiem funkcji. Oznacza to, że musisz wymyślić rodzaje cech konia – takie jak kształt, kopyta, kolor i wzrost – które algorytmy spróbują następnie zidentyfikować. Znowu jest to dobre podejście, ale daleko mu do ideału. Co się stanie, jeśli Twoje cechy są nieodpowiednie lub nie uwzględniają wartości odstających lub wyjątków? W takich przypadkach prawdopodobnie ucierpi dokładność modelu. W końcu istnieje wiele odmian konia. Wyodrębnianie funkcji ma również wadę polegającą na ignorowaniu dużej ilości danych. Może to być niezwykle skomplikowane – jeśli nie niemożliwe – w niektórych przypadkach użycia. Spójrz na wirusy

komputerowe. Ich struktury i wzory, znane jako sygnatury, nieustannie się zmieniają, aby przenikać do systemów. Ale przy ekstrakcji cech osoba musiałaby jakoś to przewidzieć, co nie jest praktyczne. To dlatego oprogramowanie cyberbezpieczeństwa często zbiera sygnatury po tym, jak wirus wyrządził szkody. Ale dzięki głębokiemu uczeniu możemy rozwiązać te problemy. Podejście to analizuje wszystkie dane piksel po pikselu, a następnie odnajduje zależności za pomocą sieci neuronowej, która naśladuje ludzki mózg. Spójrzmy...

Czym więc jest głębokie uczenie?

Głębokie uczenie to poddziedzina uczenia maszynowego. Ten rodzaj systemu pozwala na przetwarzanie ogromnych ilości danych w celu znalezienia relacji i wzorców, których ludzie często nie są w stanie wykryć. Słowo „głębokie” odnosi się do liczby ukrytych warstw w sieci neuronowej, które zapewniają dużą moc uczenia się. Jeśli chodzi o sztuczną inteligencję, głębokie uczenie jest w czołówce i często generuje większość szumu w mediach głównego nurtu. „[Głębokie uczenie] AI to nowa elektryczność” – wychwalał Andrew Yan-Tak Ng, były główny naukowiec w Baidu i współzałożyciel Google Brain. Ale ważne jest również, aby pamiętać, że uczenie głębokie jest wciąż na wczesnym etapie rozwoju i komercjalizacji. Na przykład dopiero około 2015 roku Google zaczął używać tej technologii w swojej wyszukiwarce. Jak widzieliśmy w rozdziale 1, historia sieci neuronowych była pełna przyptyków i przepłyłów. To Frank Rosenblatt stworzył perceptron, który był dość podstawowym systemem. Jednak prawdziwy postęp akademicki w zakresie sieci neuronowych nastąpił dopiero w latach 80., tak jak w przypadku przełomów w zakresie wstecznej propagacji, spłotowych sieci neuronowych i sieci neuronowych rekurencyjnych. Aby jednak uczenie głębokie miało wpływ na rzeczywisty świat, potrzebny byłby oszałamiający wzrost ilości danych, na przykład z Internetu, oraz gwałtowny wzrost mocy obliczeniowej.

Mózg i głębokie uczenie

Ważący zaledwie około 3,3 funta ludzki mózg jest niesamowitym wyczynem ewolucji. Istnieje około 86 miliardów neuronów - często nazywanych istotą szarą – które są połączone z bilionami synaps. Pomyśl o neuronach jako procesorach (Central Processing Units), które pobierają dane. Uczenie następuje wraz ze wzmocnieniem lub osłabieniem synaps. Mózg składa się z trzech regionów: przodomózgowia, śródmózgowia i tyłomózgowia. Wśród nich jest wiele obszarów, które pełnią różne funkcje. Niektóre z głównych to:

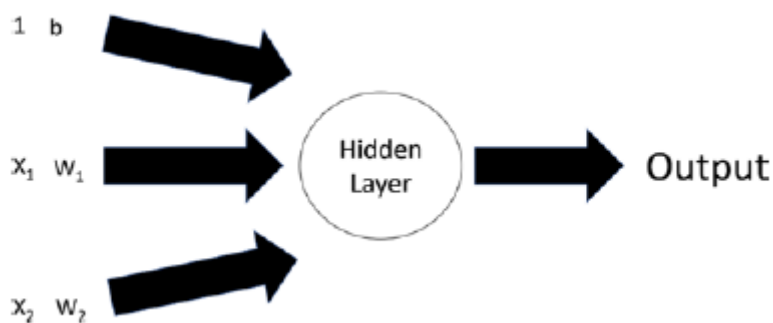
- **Hipokamp:** To tutaj twój mózg przechowuje wspomnienia. W rzeczywistości jest to część, która zawodzi, gdy osoba cierpi na chorobę Alzheimera, w której osoba traci zdolność do tworzenia wspomnień krótkotrwałych.
- **Płat czołowy:** tutaj mózg skupia się na emocjach, mowie, kreatywności, ocenie, planowaniu i rozumowaniu.
- **Kora mózgowa:** jest to prawdopodobnie najważniejsza cecha, jeśli chodzi o sztuczną inteligencję. Kora mózgowa pomaga w myśleniu i innych czynnościach poznawczych. Według badań Suzany Herculano-Houzel poziom inteligencji jest powiązany z liczbą neuronów w tym obszarze mózgu.

Zatem jak głębokie uczenie się wypada w porównaniu z ludzkim mózgiem? Istnieją pewne wątpliwości. Przynajmniej w obszarach takich jak siatkówka istnieje proces pozyskiwania danych i przetwarzania ich przez złożoną sieć, która opiera się na przypisywaniu wag. Ale oczywiście jest to tylko minutowa część procesu uczenia się. Poza tym ludzki mózg wciąż kryje wiele tajemnic i oczywiście nie opiera się on na takich rzeczach jak komputery cyfrowe (zamiast tego wydaje się, że jest to raczej

system analogowy). Jednak wraz z postępem badań odkrycia w neuronauce mogą pomóc w tworzeniu nowych modeli sztucznej inteligencji.

Sztuczne sieci neuronowe (SSN)

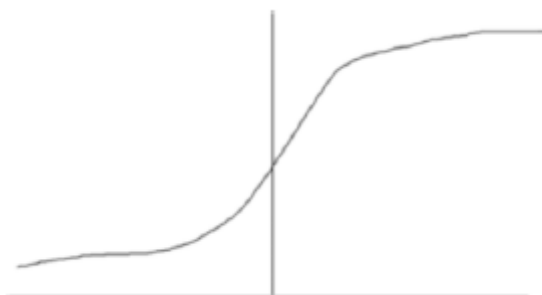
Na najbardziej podstawowym poziomie sztuczna sieć neuronowa (SSN) to funkcja obejmująca jednostki (które mogą być również nazywane neuronami, perceptronami lub węzłami). Każda jednostka będzie miała wartość i wagę, które wskazują na względną wagę, i trafi do ukrytej warstwy. Warstwa ukryta używa funkcji, a wynik staje się wynikiem. Istnieje również inna wartość, zwana bias, która jest stałą i jest używana do obliczania funkcji. Ten rodzaj uczenia modelu nazywa się siecią neuronową ze sprzężeniem do przodu. Innymi słowy, przechodzi tylko od wejścia do ukrytej warstwy do wyjścia. Nie wraca. Ale może trafić do nowej sieci neuronowej, w której dane wyjściowe stają się danymi wejściowymi. Rysunek przedstawia wykres sieci neuronowej ze sprzężeniem do przodu.



Zagłębnym się w to, biorąc przykład. Załóżmy, że tworzysz model, aby przewidzieć, czy akcje firmy wzrosną. Poniżej przedstawiono, co reprezentują zmienne, a także przypisane wartości i wagi:

- X1: Przychody rosną o co najmniej 20% rocznie. Wartość to 2.
- X2: Marża zysku wynosi co najmniej 20%. Wartość to 4.
- W1: 1.9.
- W2: 9.6.
- b: To jest odchylenie (wartość wynosi 1), które pomaga wygładzić obliczenia.

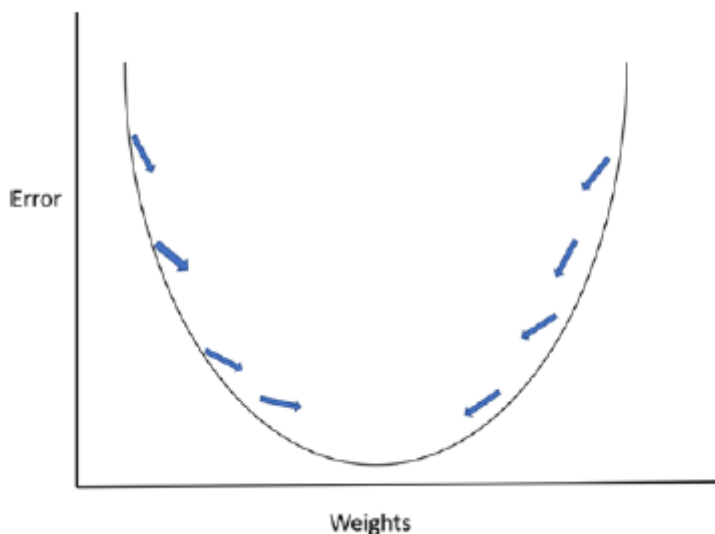
Następnie zsumujesz wagi, a następnie funkcja przetworzy informacje. Często wiąże się to z funkcją aktywacji, która jest nieliniowa. Odzwierciedla to bardziej rzeczywisty świat, ponieważ dane zwykle nie są w linii prostej. Teraz do wyboru jest wiele funkcji aktywacji. Jednym z najczęstszych jest esicy. Powoduje to kompresję wartości wejściowej do zakresu 0–1. Im bliżej 1, tym dokładniejszy model. Kiedy narysujesz tę funkcję, będzie ona wyglądać jak kształt litery S. Patrz rysunek



Jak widać, system jest stosunkowo uproszczony i nie będzie pomocny w zaawansowanych modelach AI. Aby dodać znacznie więcej mocy, zwykle potrzeba wielu ukrytych warstw. Powoduje to wielowarstwowy perceptron (MLP). Pomaga również użyć czegoś, co nazywa się propagacją wsteczną, która pozwala na zakręcenie wyjścia z powrotem do sieci neuronowej.

Propagacja wsteczna

Jedną z głównych wad sztucznych sieci neuronowych jest proces wprowadzania korekt wag w modelu. Tradycyjne podejścia, takie jak zastosowanie algorytmu mutacji, wykorzystywały wartości losowe, które okazały się czasochłonne. Biorąc to pod uwagę, naukowcy szukali alternatyw, takich jak wsteczna propagacja. Ta technika istniała od lat 70., ale nie cieszyła się dużym zainteresowaniem, ponieważ brakowało jej wydajności. Ale David Rumelhart, Geoffrey Hinton i Ronald Williams zdali sobie sprawę, że wsteczna propagacja wciąż ma potencjał, o ile została udoskonalona. W 1986 r. napisali artykuł zatytułowany „Uczenie się reprezentacji przez błędy wstecznej propagacji błędów” i był to bomba w społeczności AI. Wyraźnie pokazało, że wsteczna propagacja może być znacznie szybsza, ale także umożliwia tworzenie potężniejszych sztucznych sieci neuronowych. Nie powinno dziwić, że w propagację wsteczną jest dużo matematyki. Ale kiedy się sprowadza, chodzi o dostosowanie sieci neuronowej, gdy zostaną znalezione błędy, a następnie iterowanie nowych wartości przez sieć neuronową ponownie. Zasadniczo proces obejmuje niewielkie zmiany, które nadal optymalizują model. Na przykład założmy, że jedno z wejść ma wyjście 0,6. Oznacza to, że błąd wynosi 0,4 (1,0 minus 0,6), co jest subpar. Ale możemy wtedy dokonać wstecznej propagacji danych wyjściowych i być może nowe dane wyjściowe mogą osiągnąć 0,65. To szkolenie będzie trwało do momentu, gdy wartość będzie znacznie zbliżona do 1. Rysunek 4-3 ilustruje ten proces.



Na początku występuje wysoki poziom błędów, ponieważ wagi są zbyt duże. Ale robiąc iteracje, błędy będą stopniowo spadać. Jednak robienie tego zbyt wiele może oznaczać wzrost liczby błędów. Innymi słowy, celem propagacji wstecznej jest znalezienie punktu środkowego. Miarą powodzenia wstecznej propagacji było pojawienie się niezliczonych zastosowań komercyjnych. Jeden nazywał się NETtalk, który został opracowany przez Terrence'a Sejnowskiego i Charlesa Rosenberga w połowie lat 80-tych. Maszyna była w stanie nauczyć się wymawiać tekst w języku angielskim. NETtalk był tak interesujący, że został nawet zaprezentowany w programie Today. Powstało również wiele startupów wykorzystujących wsteczną propagację, takich jak HNC Software. Zbudował modele wykrywające oszustwa związane z kartami kredytowymi. Aż do tego momentu — kiedy pod koniec lat 80. założono

HNC — proces ten był wykonywany głównie ręcznie, co prowadziło do kosztownych błędów i niskich wolumenów emisji. Jednak stosując metody głębokiego uczenia się, firmy obsługujące karty kredytowe były w stanie zaoszczędzić miliardy dolarów. W 2002 roku firma HNC została przejęta przez Fair, Isaac i wyceniona na 810 milionów dolarów.

Różne sieci neuronowe

Najbardziej podstawowym typem sieci neuronowej jest w pełni połączona sieć neuronowa. Jak sama nazwa wskazuje, jest to miejsce, w którym wszystkie neurony mają połączenia między warstwami. Ta sieć jest w rzeczywistości dość popularna, ponieważ oznacza to, że przy tworzeniu modelu trzeba używać niewielkiego osądu. A jakie są inne sieci neuronowe? Typowe z nich obejmują rekurencyjną sieć neuronową (RNN), splotową sieć neuronową (CNN) i generatywną sieć kontrykcyjną (GAN), którą omówimy dalej.

Rekurencyjna sieć neuronowa

W przypadku rekurencyjnej sieci neuronowej (RNN) funkcja nie tylko przetwarza dane wejściowe, ale także wcześniejsze dane wejściowe w czasie. Przykładem tego jest to, co dzieje się, gdy wprowadzasz znaki w aplikacji do przesyłania wiadomości. Gdy zaczniesz pisać, system przewidzi słowa. Jeśli więc naciśniesz „On”, komputer zasugeruje „On”, „Cześć” i „Oto”. RNN jest zasadniczo ciągiem sieci neuronowych, które żywią się sobą na podstawie złożonych algorytmów. Istnieją odmiany modelu. Jeden nazywa się LSTM, co oznacza długą pamięć krótkotrwałą. Wynikało to z artykułu napisanego przez profesorów Seppa Hochreitera i Jürgena Schmidhubera w 1997 roku⁶. Przedstawili w nim sposób na efektywne wykorzystanie danych wejściowych, które są od siebie oddzielone przez długi czas, co pozwala na użycie większej liczby zbiorów danych. Oczywiście RNN mają wady. Występuje problem znikającego gradientu, co oznacza, że dokładność spada wraz ze wzrostem modeli. Trening modeli może zająć więcej czasu. Aby sobie z tym poradzić, Google opracował nowy model o nazwie Transformer, który jest znacznie bardziej wydajny, ponieważ równoległe przetwarza dane wejściowe. Daje to również dokładniejsze wyniki. Google uzyskał wiele informacji na temat RNN dzięki aplikacji Tłumacz, która obsługuje ponad 100 języków i przetwarza ponad 100 miliardów słów dziennie. Uruchomiony w 2006 roku, początkowo wykorzystywał systemy uczenia maszynowego. Ale w 2016 r. Google przeszło na głębokie uczenie, tworząc Google Neural Machine Translation.⁸ Podsumowując, zaowocowało to znacznie wyższymi wskaźnikami dokładności. Zastanów się, jak Tłumacz Google pomógł lekarzom, którzy pracują z pacjentami posługującymi się innymi językami. Według badania przeprowadzonego na Uniwersytecie Kalifornijskim w San Francisco (UCSF), opublikowanego w JAMA Internal Medicine, aplikacja miała wskaźnik dokładności na poziomie 92% z tłumaczeniami z języka angielskiego na hiszpański. To wzrosło z 60% w ciągu ostatnich kilku lat.

Konwolucyjna sieć neuronowa (CNN)

Intuicyjnie, sensowne jest, aby wszystkie jednostki w sieci neuronowej były połączone. Działa to dobrze z wieloma aplikacjami. Ale są scenariusze, w których jest to dalekie od optymalnego, takie jak rozpoznawanie obrazu. Wyobraź sobie, jak skomplikowany byłby model, w którym każdy piksel jest jednostką! Może szybko stać się nie do opanowania. Pojawiłyby się również inne komplikacje, takie jak nadmierne dopasowanie. W tym miejscu dane nie odzwierciedlają tego, co jest testowane lub skupiają się na niewłaściwych funkcjach. Aby sobie z tym wszystkim poradzić, możesz użyć splotowej sieci neuronowej (CNN). Początki tego zjawiska sięgają profesora Yanna LeCuna w 1998 roku, kiedy opublikował artykuł zatytułowany „Gradient-Based Learning Applied to Document Recognition”¹¹. Jednak gdy w 2012 r. uczenie głębokie zaczęło wykazywać znaczne postępy, naukowcy ponownie przyjrzeni się modelowi. LeCun czerpał inspirację dla CNN od zdobywców Nagrody Nobla Davida Hubela i Torstena Wiesela, którzy badali neurony kory wzrokowej. System ten pobiera obraz z siatkówki i

przetwarza go na różnych etapach – od łatwego do bardziej złożonego. Każdy z etapów nazywany jest splotem. Na przykład pierwszym poziomem byłoby zidentyfikowanie linii i kątów; następnie kora wzrokowa znajdzie kształty; a następnie wykryje obiekty. Jest to analogiczne do tego, jak działa komputerowa CNN. Weźmy przykład: założmy, że chcesz zbudować model, który może identyfikować literę. CNN będzie mieć dane wejściowe w postaci obrazu, który ma 3072 piksele. Każdy z pikseli będzie miał wartość od 0 do 255, co wskazuje na ogólną intensywność. Korzystając z CNN, komputer przejdzie przez wiele odmian, aby zidentyfikować cechy. Pierwsza to warstwa splotowa, czyli filtr skanujący obraz. W naszym przykładzie może to być 5×5 pikseli. Proces utworzy mapę funkcji, która jest długą tablicą liczb. Następnie model zastosuje więcej filtrów do obrazu. W ten sposób CNN zidentyfikuje linie, krawędzie i kształty – wszystkie wyrażone w liczbach. W przypadku różnych warstw wyjściowych model użyje puli, która łączy je w celu wygenerowania jednego wyniku, a następnie stworzenia w pełni połączonej sieci neuronowej. CNN z pewnością może się skomplikować. Powinien jednak być w stanie dokładnie zidentyfikować liczby wprowadzane do systemu.

Generatywne sieci kontryktoryjne (GAN)

Ian Goodfellow, który uzyskał tytuł magistra informatyki na Stanford i doktorat z uczenia maszynowego na Université de Montréal, zaczął pracować w Google. W wieku 20 lat był współautorem jednej z najlepszych książek o sztucznej inteligencji, zatytułowanej Deep Learning, a także wprowadzał innowacje dzięki Google Maps. Ale to właśnie w 2014 roku miał swój największy przełom. Tak naprawdę zdarzyło się to w pubie w Montrealu, kiedy rozmawiał z niektórymi ze swoich przyjaciół o tym, jak głębokie uczenie może tworzyć zdjęcia. W tamtym czasie podejściem było stosowanie modeli generatywnych, ale często były one rozmyte i bezsensowne.

Goodfellow zdał sobie sprawę, że musi być lepsze dłaczego. Dłaczego więc nie wykorzystać teorii gier? Oznacza to, że dwa modele rywalizują ze sobą w ściślejszej pętli sprzężenia zwrotnego. Można to również zrobić z danymi nieoznakowanymi. Oto podstawowy przepływ pracy:

- Generator: ta sieć neuronowa tworzy mnóstwo nowych kreacji, takich jak zdjęcia lub zdania.
- Dyskryminator: ta sieć neuronowa analizuje kreacje, aby zobaczyć, które z nich są prawdziwe.
- Korekty: przy tych dwóch wynikach nowy model zmieniłby kreacje, aby były jak najbardziej realistyczne.

Dzięki wielu powtórzeniom dyskryminator nie będzie już musiał być używany. Był tak podekscytowany pomysłem, że po wyjściu z pubu zaczął kodować swoje pomysły. W rezultacie powstał nowy model głębokiego uczenia: generatywna sieć adwersarzy lub GAN. A wyniki były wyjątkowe. Wkrótce miał zostać gwiazdą rocka AI. Badania GAN zaowocowały już ponad 500 artykułami akademickimi. Firmy takie jak Facebook również wykorzystały tę technologię, na przykład do analizy i przetwarzania zdjęć. Główny naukowiec firmy zajmujący się sztuczną inteligencją, Yann LeCun, zauważył, że GAN to „najfajniejszy pomysł na głębokie uczenie w ciągu ostatnich 20 lat”. Wykazano również, że GAN pomagają w zaawansowanych badaniach naukowych. Na przykład pomogli poprawić dokładność wykrywania zachowania cząstek subatomowych w Wielkim Zderzaczem Hadronów w CERN w Szwajcarii. Jeszcze we wczesnych fazach ta technologia może prowadzić do takich rzeczy, jak komputer, który może opracowywać nowe rodzaje modnych przedmiotów a może nowomodne urządzenie do noszenia. Być może GAN mógłby nawet wymyślić hitową piosenkę rapową. I może być wcześniej niż myślisz. Jako nastolatek Robbie Barrat nauczył się korzystać z systemów głębokiego uczenia i zbudował model do rapowania w stylu Kanye Westa. Ale to był dopiero początek jego magii AI. Jako badacz w Stanford opracował własną platformę GAN, która przetworzyła około 10 000 nagich portretów. System tworzyłby wtedy naprawdę fascynujące nowe dzieła sztuki (można je znaleźć na jego koncie na

Twitterze @DrBeef_). Aha, a także stworzył swój system jako open source na swoim koncie GitHub. Zwróciło to uwagę kolektywu francuskich artystów o nazwie Obvious, który wykorzystał tę technologię do stworzenia portretów osiemnastowiecznej fikcyjnej rodziny. Polegała na przetworzeniu 15 000 portretów z XIV-XX wieku. W 2018 roku Obvious wystawił swoje dzieło na aukcji Christie's, przynosząc 432 000 \$. Ale niestety, jeśli chodzi o GAN, zdarzały się zastosowania, które były mniej niż godne podziwu. Jednym z przykładów jest użycie ich do deepfake, które polega na wykorzystaniu sieci neuronowych do tworzenia obrazów lub filmów, które wprowadzają w błąd. Niektóre z nich są po prostu zabawne. Na przykład jeden GAN pozwala Barackowi Obamie powiedzieć wszystko, co mu powiesz! Jednak istnieje wiele zagrożeń. Naukowcy z New York University i Michigan State University napisali artykuł, który skupił się na „DeepMasterPrints. Pokazał, jak GAN może opracować fałszywe odciski palców, aby odblokować trzy rodzaje smartfonów! Potem był incydent z tak zwanym deepfake wideo aktorki Jennifer Lawrence na konferencji prasowej Złoty Globów. Jej twarz połączyła się z twarzą Steve'a Buscemi

Aplikacje do głębokiego uczenia

Przy tak dużej ilości pieniędzy i zasobów przeznaczanych na głębokie uczenie nastąpił gwałtowny wzrost innowacji. Wydaje się, że każdego dnia ogłasza się coś niesamowitego. Więc jakie są niektóre aplikacje? Gdzie okazało się, że głębokie uczenie się zmienia zasady gry? Rzućmy okiem na niektóre, które obejmują obszary takie jak opieka zdrowotna, energia, a nawet trzęsienia ziemi

Przypadek użycia: wykrywanie choroby Alzheimera

Pomimo dziesięcioleci badań, lekarstwo na chorobę Alzheimera pozostaje nieuchwytnie. Chociaż naukowcy opracowali leki, które spowolniły postęp choroby. W związku z tym wczesna diagnoza ma kluczowe znaczenie, a głębokie uczenie się może być potencjalnie dużą pomocą. Naukowcy z Zakładu Radiologii i Biomedycyny UCSF

Obrazowanie wykorzystało tę technologię do analizy ekranów mózgu - z publicznego zbioru danych inicjatywy Alzheimer's Disease Neuroimaging Initiative - i do wykrywania zmian w poziomie glukozy. Wynik: model może diagnozować chorobę Alzheimera do sześciu lat przed diagnozą kliniczną. Jeden z testów wykazał 92% dokładność, a drugi 98%. Teraz jest to wciąż w początkowej fazie — i będzie trzeba przeanalizować więcej zbiorów danych. Ale jak dotąd wyniki są bardzo zachęcające. Według dr Jae Ho Sohn, autorki badania: Jest to idealne zastosowanie uczenia głębokiego, ponieważ jest ono szczególnie silne w wykrywaniu bardzo subtelnych, ale rozproszonych procesów. Radiolodzy są naprawdę silni w identyfikowaniu niewielkich ogniskowych, takich jak guz mózgu, ale walczymy z wykrywaniem wolniejszych, globalnych zmian. Biorąc pod uwagę siłę głębokiego uczenia się w tego typu aplikacjach, zwłaszcza w porównaniu z ludźmi, wydawało się to naturalnym zastosowaniem.

Zastosowanie: energia

Ze względu na ogromną infrastrukturę centrum danych Google jest jednym z największych konsumentów energii. Nawet niewielka poprawa wydajności może mieć znaczny wpływ na wynik finansowy. Ale mogą też być korzyści z mniejszej emisji dwutlenku węgla. Aby pomóc w realizacji tych celów, jednostka DeepMind firmy Google stosuje głębokie uczenie, które obejmuje lepsze zarządzanie energią wiatrową. Mimo że jest to czyste źródło energii, może być trudne w użyciu ze względu na zmiany pogody. Ale algorytmy głębokiego uczenia DeepMind były krytyczne. Przy zastosowaniu 700 megawatów energii wiatrowej w Stanach Zjednoczonych byli w stanie sporządzić dokładne prognozy dotyczące produkcji z czasem realizacji wynoszącym 36 godzin. Według bloga DeepMind: Jest to ważne, ponieważ źródła energii, które można zaplanować (tj. mogą dostarczyć określoną ilość energii elektrycznej w ustalonym czasie) są często bardziej wartościowe dla sieci... Do tej pory uczenie

maszynowe zwiększyło wartość naszego wiatru energii o około 20 procent w porównaniu ze scenariuszem bazowym, w którym brak zobowiązań czasowych do sieci. Ale oczywiście ten system głębokiego uczenia może być czymś więcej niż tylko Google – może mieć szeroki wpływ na zużycie energii na całym świecie.

Zastosowanie: trzęsienia ziemi

Trzęsienia ziemi są niezwykle trudne do zrozumienia. Są też niezwykle trudne do przewidzenia. Musisz ocenić uskoki, formacje skalne i deformacje, aktywność elektromagnetyczną i zmiany w wodzie gruntowej. Hej, są nawet dowody na to, że zwierzęta potrafią wyczuć trzęsienie ziemi! Ale przez dziesięciolecia naukowcy zebrali ogromne ilości danych na ten temat. Innymi słowy, może to być aplikacja do głębokiego uczenia się, prawda? Absolutnie. Sejsmolodzy z Caltech, do których należą Yisong Yue, Egill Hauksson, Zachary Ross i Men-Andrin Meier, prowadzą znaczne badania na ten temat, używając splotowych i rekurencyjnych sieci neuronowych. Próbuje zbudować skuteczny system wczesnego ostrzegania. Oto, co miał do powiedzenia Yue:

Sztuczna inteligencja może [analizować trzęsienia ziemi] szybciej i dokładniej niż ludzie, a nawet znajdować wzorce, które w przeciwnym razie umknęłyby ludzkiemu oku. Co więcej, wzorce, które mamy nadzieję wyodrębnić, są trudne do odpowiedniego uchwycenia przez systemy oparte na regułach, a zatem zaawansowane możliwości dopasowywania wzorców współczesnego głębokiego uczenia mogą zapewnić lepszą wydajność niż istniejące zautomatyzowane algorytmy monitorowania trzęsień ziemi.

Ale kluczem jest poprawa gromadzenia danych. Oznacza to więcej analiz małych trzęsień ziemi (w Kalifornii jest ich średnio 50 dziennie). Celem jest stworzenie katalogu trzęsień ziemi, który może doprowadzić do stworzenia wirtualnego sejsmologa, który może dokonać oceny trzęsienia ziemi szybciej niż człowiek. Może to pozwolić na skrócenie czasu realizacji w przypadku trzęsienia ziemi, co może pomóc uratować życie i mienie.

Przypadek użycia: Radiologia

Skany PET i MRI to niesamowita technologia. Ale są zdecydowanie wady. Pacjent musi pozostać w rurce ograniczającej przez 30 minut do godziny. Jest to niewygodne i oznacza narażenie na gadolin, który, jak wykazano, ma szkodliwe skutki uboczne. Greg Zaharchuk i Enhao Gong, którzy spotkali się w Stanford, myśleli, że może być lepszy sposób. Zaharchuk był doktorem medycyny i doktorem ze specjalizacją z radiologii. Był również promotorem doktoratu Gong, który był doktorem elektrotechniki w zakresie głębokiego uczenia się i rekonstrukcji obrazu medycznego. W 2017 roku założyli firmę Subtle Medical i zatrudnili jednych z najzdolniejszych naukowców zajmujących się obrazowaniem, radiologów i ekspertów AI. Wspólnie podjęli wyzwanie ulepszenia skanów PET i MRI. Firma Subtle Medical stworzyła system, który nie tylko skraca czas skanowania MRI i PET nawet dziesięciokrotnie, ale dokładność jest znacznie wyższa. To było zasilane przez wysokiej klasy procesory graficzne NVIDIA. Następnie w grudniu 2018 r. system otrzymał dopuszczenie FDA (Federal Drug Administration) 510(k) oraz dopuszczenie znaku CE na rynek europejski. Było to pierwsze w historii jądrowe urządzenie medyczne oparte na sztucznej inteligencji, które uzyskało oba te oznaczenia. Subtle Medical ma więcej planów, aby zrewolucjonizować biznes radiologiczny. Od 2019 roku opracowuje SubtleMRM, który będzie jeszcze silniejszy niż obecne rozwiązanie firmy, oraz SubtleGADTM, który zmniejszy dawki gadolinu

Sprzęt do głębokiego uczenia

Jeśli chodzi o układy scalone do głębokiego uczenia się, głównym wyborem były procesory graficzne. Ale w miarę jak sztuczna inteligencja staje się coraz bardziej wyrafinowana - na przykład w przypadku sieci GAN - i zbiory danych są znacznie większe, z pewnością jest więcej miejsca na nowe podejścia. Firmy mają również niestandardowe potrzeby, takie jak funkcje i dane. W końcu aplikacja dla konsumenta jest zwykle zupełnie inna niż aplikacja skoncentrowana na przedsiębiorstwie.

W rezultacie niektóre z wielkich firm technologicznych opracowały własne chipsety:

- Google: Latem 2018 r. firma ogłosiła trzecią wersję swojego procesora Tensor Processing Unit (TPU; pierwszy układ został opracowany w 2016 r.) do chłodzenia cieczą w centrach danych. Google ogłosił również wersję swojego TPU dla urządzeń. Zasadniczo oznacza to, że przetwarzanie będzie miało mniejsze opóźnienia, ponieważ nie będzie potrzeby uzyskiwania dostępu do chmury.
- Amazon: w 2018 roku firma ogłosiła AWS Inferentia.²⁶ Technologia, która powstała w wyniku przejścia Annapurny w 2015 roku, koncentruje się na obsłudze złożonych operacji wnioskowania. Innymi słowy, dzieje się tak po wytrenowaniu modelu.
- Facebook i Intel: Firmy te połączyły siły, aby stworzyć chip AI.²⁷ Jednak inicjatywa jest wciąż w początkowej fazie. Intel również zyskuje przyczepność dzięki chipowi AI o nazwie Nervana Neural Network Processor (NNP).
- Alibaba: Firma stworzyła własną firmę produkującą układy sztucznej inteligencji o nazwie Pingtoug.²⁸ Ma również plany zbudowania procesora komputera kwantowego opartego na kubitach (reprezentują cząstki subatomowe, takie jak elektrony i fotony).
- Tesla: Elon Musk opracował swój własny chip AI. Ma 6 miliardów tranzystorów i może przetwarzać 36 bilionów operacji na sekundę

Istnieje wiele startupów, które również grają na rynku chipów AI. Wśród wiodących firm znajduje się Untether AI, która koncentruje się na tworzeniu chipów zwiększających prędkości przesyłania danych (jest to szczególnie trudna część AI). W jednym z prototypów firmy proces ten był o ponad 1000 szybszy niż w przypadku typowego układu AI.³⁰ Intel wraz z innymi inwestorami wziął udział w rundzie finansowania o wartości 13 milionów dolarów w 2019 r. Teraz, jeśli chodzi o układy AI, NVIDIA ma dominujący udział w rynku. Jednak ze względu na znaczenie tej technologii wydaje się nieuniknione, że na rynku pojawi się coraz więcej ofert.

Kiedy używać głębokiego uczenia?

Ze względu na moc głębokiego uczenia istnieje pokusa, aby najpierw użyć tej technologii podczas tworzenia projektu AI. Ale to może być duży błąd. Głębokie uczenie nadal ma wąskie przypadki użycia, takie jak w przypadku zestawów danych tekstowych, wideo, graficznych i szeregów czasowych. Istnieje również zapotrzebowanie na duże ilości danych i systemy komputerowe o dużej mocy. Aha, a głębokie uczenie się jest lepsze, gdy wyniki można określić ilościowo i zweryfikować. Aby zobaczyć dlaczego, rozważmy następujący przykład. Zespół naukowców kierowany przez Thomasa Hartunga (toksykologa z Johns Hopkins University) stworzył zbiór danych zawierający około 10 000 substancji chemicznych, które zostały oparte na 800 000 testów na zwierzętach. Dzięki wykorzystaniu głębokiego uczenia wyniki wykazały, że model był bardziej predykcyjny niż wiele testów na zwierzętach pod kątem toksyczności. Pamiętaj, że testy na zwierzętach mogą być nie tylko kosztowne i wymagać środków bezpieczeństwa, ale także dawać niespójne wyniki z powodu powtarzanych testów na tej samej substancji chemicznej. „Pierwszy scenariusz ilustruje moc predykcyjną głębokiego uczenia i jego zdolność do odkrywania korelacji z dużych zbiorów danych, których człowiek nigdy by nie znalazł” - powiedział Sheldon Fernandez, który jest dyrektorem generalnym DarwinAI. Więc gdzie jest scenariusz, w którym głębokie

uczenie się nie sprawdza? W rzeczywistości ilustracją tego są Mistrzostwa Świata FIFA 2018 w Rosji, które wygrała Francja. Wielu badaczy próbowało przewidzieć wyniki wszystkich 64 meczów, ale wyniki były dalekie od dokładności:

- Jedna grupa badaczy zastosowała model konsensusu bukmacherskiego, który wskazywał, że Brazylia wygra.
- Inna grupa badaczy wykorzystała algorytmy, takie jak losowy las i ranking Poissona, aby przewidzieć, że Hiszpania zwycięży.

Problem polega na tym, że trudno jest znaleźć odpowiednie zmienne, które mają moc predykcyjną. W rzeczywistości modele uczenia głębokiego w zasadzie nie są w stanie poradzić sobie ze złożonością funkcji niektórych wydarzeń, zwłaszcza tych, które mają elementy chaotyczne. Jednak nawet jeśli dysponujesz odpowiednią ilością danych i mocą obliczeniową, nadal musisz zatrudniać osoby z doświadczeniem w głębokim uczeniu, co nie jest łatwe. Należy pamiętać, że wybór odpowiedniego modelu i dopracowanie go to nie lada wyzwanie. Ile powinno być hiperparametrów? Jaka powinna być liczba ukrytych warstw? A jak oceniasz model? Wszystko to jest bardzo złożone. Nawet eksperci mogą się pomylić. Oto następująca wiadomość od Sheldona: Jeden z naszych klientów z branży motoryzacyjnej spotkał się z dziwnym zachowaniem, w którym autonomiczny samochód coraz częściej skręcał w lewo, gdy niebo miało określony odcień fioletu. Po miesiącach bolesnego debugowania ustalili, że szkolenie dotyczące pewnych scenariuszy zwrotnych zostało przeprowadzone na pustyni Nevada, gdy niebo miało szczególny odcień. Bez wiedzy ludzkich projektantów sieć neuronowa ustaliła korelację między jej zachowaniem obrotowym a niebiańskim odcieniem. Istnieje kilka narzędzi, które pomagają w procesie głębokiego uczenia się, takich jak SageMaker firmy Amazon.com, HyperTune firmy Google i SigOpt. Ale przed nami jeszcze długa droga. Jeśli uczenie głębokie nie pasuje, możesz rozważyć uczenie maszynowe, które często wymaga stosunkowo mniej danych. Co więcej, modele wydają się być znacznie prostsze, ale wyniki mogą być nadal bardziej efektywne.

Wady Deep Learning

Biorąc pod uwagę wszystkie innowacje i przełomy, rozsądne jest, że wiele osób uważa głębokie uczenie się za srebrną kulę. Oznacza to, że nie będziemy już musieli prowadzić samochodu. Może nawet oznaczać, że wyleczymy raka. Jak nie być podekscytowanym i optymistycznym? To naturalne i rozsądne. Należy jednak zauważyć, że głębokie uczenie jest wciąż w początkowej fazie i faktycznie istnieje wiele dokuczliwych problemów. Dobrym pomysłem jest złagodzenie oczekiwań. W 2018 roku Gary Marcus napisał artykuł zatytułowany „Głębokie uczenie: krytyczna ocena”, w którym jasno przedstawił wyzwania. W swoim artykule zauważa:

Na tle znacznego postępu w takich obszarach, jak rozpoznawanie mowy, rozpoznawanie obrazów i granie w gry oraz znacznego entuzjazmu w prasie popularnej, przedstawiam dziesięć obaw dotyczących uczenia głębokiego i sugeruję, że uczenie głębokie musi być uzupełnione innymi technikami, jeśli aby osiągnąć sztuczną ogólną inteligencję. Marcus z pewnością ma odpowiedni rodowód, aby przedstawić swoje obawy, ponieważ ma zarówno akademickie, jak i biznesowe doświadczenie w sztucznej inteligencji. Zanim został profesorem na Wydziale Psychologii Uniwersytetu Nowojorskiego, sprzedał swój startup o nazwie Geometric Intelligence firmie Uber. Marcus jest także autorem kilku bestsellerowych książek, takie jak *The Haphazard Construction of the Human Mind*. Oto kilka z jego obaw związanych z głębokim uczeniem:

- Czarna skrzynka: Model głębokiego uczenia może z łatwością zawierać miliony parametrów, które obejmują wiele ukrytych warstw. Jasne zrozumienie tego jest naprawdę poza możliwościami danej osoby. To prawda, że niekoniecznie musi to być problem z rozpoznawaniem kotów w zbiorze danych.

Ale z pewnością może to być problem z modelami do diagnostyki medycznej lub określania bezpieczeństwa platformy wiertniczej. W takich sytuacjach regulatorzy będą chcieli dobrze zrozumieć przejrzystość modeli. Z tego powodu naukowcy szukają systemów do określania „wyjaśnialności”, co zapewnia zrozumienie modeli głębokiego uczenia się.

- Dane: ludzki mózg ma swoje wady. Ale są pewne funkcje, które bardzo dobrze przypominają zdolność uczenia się przez abstrakcję. Załóżmy na przykład, że Jan, która ma pięć lat, idzie z rodziną do restauracji. Jej matka wskazuje przedmiot na talerzu i mówi, że to „taco”. Nie musi tego wyjaśniać ani podawać żadnych informacji na ten temat. Zamiast tego mózg Jana natychmiast przetworzy te informacje i zrozumie ogólny wzór. W przyszłości, gdy zobaczy kolejne taco – nawet jeśli ma różnice, na przykład z ubiorem – będzie wiedziała, co to jest. W większości jest to intuicyjne. Ale niestety, jeśli chodzi o głębokie uczenie się, nie ma nauki taco przez abstrakcję! System musi przetworzyć ogromne ilości informacji, aby je rozpoznać. Oczywiście nie stanowi to problemu dla firm takich jak Facebook, Google, czy nawet Uber. Jednak wiele firm ma znacznie bardziej ograniczone zbiory danych. W rezultacie głębokie uczenie się może nie być dobrą opcją.

- Struktura hierarchiczna: Ten sposób organizowania nie istnieje w głębokim uczeniu się. Z tego powodu zrozumienie języka ma jeszcze długą drogę do przebycia (zwłaszcza w przypadku długich dyskusji).

- Wnioskowanie otwarte: Marcus zauważa, że uczenie głębokie nie może zrozumieć niuansów między „Jan obiecał, że Mary odejdzie” a „Jan obiecał, że opuści Mary”. Co więcej, głębokie uczenie jest dalekie od możliwości, na przykład, przeczytania Dumy i uprzedzenia Jane Austen i umiejętności odgadnięcia motywacji postaci Elizabeth Bennet.

- Myślenie koncepcyjne: Głębokie uczenie się nie pozwala na zrozumienie pojęć takich jak demokracja, sprawiedliwość czy szczęście. Nie ma też wyobraźni, wymyślenia nowych pomysłów czy planów.

- Zdrowy rozsądek: To jest coś, czego głębokie uczenie się nie sprawdza. Jeśli już, oznacza to, że model można łatwo pomylić. Na przykład, powiedzmy, że pytasz system AI: „Czy można zrobić komputer z gąbką?” W większości prawdopodobnie nie będzie wiedział, że to śmieszne pytanie.

- Przyczyna: głębokie uczenie nie jest w stanie tego określić. Wszystko sprowadza się do znajdowania korelacji.

- Wcześniejsza wiedza: CNN mogą pomóc w uzyskaniu pewnych wcześniejszych informacji, ale jest to ograniczone. Głębokie uczenie jest nadal dość samowystarczalne, ponieważ rozwiązuje tylko jeden problem na raz. Nie może pobierać danych i tworzyć algorytmów obejmujących różne domeny. Ponadto model się nie dostosowuje. W przypadku zmiany danych należy przeszkolić i przetestować nowy model. I wreszcie, uczenie głębokie nie ma wcześniejszego zrozumienia tego, co ludzie wiedzą instynktownie — na przykład podstaw fizyki rzeczywistego świata. To jest coś, co musi być wyraźnie zaprogramowane w systemie AI.

- Statyczny: głębokie uczenie działa najlepiej w dość prostych środowiskach. Właśnie dlatego sztuczna inteligencja jest tak skuteczna w grach planszowych, które mają jasny zestaw zasad i granic. Ale prawdziwy świat jest chaotyczny i nieprzewidywalny. Oznacza to, że głębokie uczenie się może spaść krótki ze złożonymi problemami, nawet z samojezdnymi samochodami.

- Zasoby: Model uczenia głębokiego często wymaga ogromnej mocy procesora, tak jak w przypadku procesorów graficznych. To może być kosztowne. Chociaż jedną z opcji jest skorzystanie z usługi w chmurze innej firmy.

To dość dużo? To prawda. Ale papier nadal pomija pewne wady. Oto kilka innych:

- Efekt motyla: Ze względu na złożoność danych, sieci i połączeń, drobna zmiana może mieć duży wpływ na wyniki modelu głębokiego uczenia się. Może to łatwo prowadzić do wniosków, które są błędne lub mylące.
- Overfitting.

Jeśli chodzi o Marcusa, jego największą obawą jest to, że sztuczna inteligencja może „zostać uwięziona w lokalnym minimum, zagnieżdżając się zbyt mocno w niewłaściwej części przestrzeni intelektualnej, zbyt mocno skupiając się na szczegółowej eksploracji określonej klasy dostępnych, ale ograniczonych modeli, które są nastawione na przechwytywanie nisko wiszących owoców – potencjalnie zaniedbując bardziej ryzykowne wycieczki, które mogą ostatecznie prowadzić do bardziej solidnej ścieżki”. Nie jest jednak pesymistą. Uważa, że naukowcy muszą wyjść poza głębokie uczenie i znaleźć nowe techniki, które mogą rozwiązać trudne problemy.

Wniosek

Chociaż Marcus wskazał na wady głębokiego uczenia się, faktem jest, że to podejście AI jest nadal niezwykle skuteczne. W niecałą dekadę zrewolucjonizował świat technologii, a także znacząco wpływa na takie obszary, jak finanse, robotyka i opieka zdrowotna. Wraz ze wzrostem inwestycji ze strony dużych firm technologicznych i VC, modele będą wprowadzać dalsze innowacje. Zachęci to również inżynierów do uzyskania stopni podyplomowych, tworząc pozytywny cykl przełomu.

Kluczowe dania na wynos

- Głębokie uczenie, które jest poddziedziną uczenia maszynowego, przetwarza ogromne ilości danych w celu wykrycia relacji i wzorców, których ludzie często nie są w stanie wykryć. Słowo „głębokie” opisuje liczbę ukrytych warstw.
- Sztuczna sieć neuronowa (ANN) to funkcja obejmująca jednostki posiadające wagi i służąca do przewidywania wartości w modelu AI.
- Ukryta warstwa jest częścią modelu, która przetwarza przychodzące dane.
- Sieć neuronowa typu feed-forward zawiera dane, które przechodzą tylko od wejścia do warstwy ukrytej do wyjścia. Wyniki nie wracają. Mogą jednak wejść do innej sieci neuronowej.
- Funkcja aktywacji jest nieliniowa. Innymi słowy, lepiej odzwierciedla rzeczywisty świat.
- Sigmoid to funkcja aktywacji, która kompresuje wartość wejściową do zakresu 0–1, co ułatwia analizę.
- Propagacja wsteczna to zaawansowana technika dostosowywania wag w sieci neuronowej. Takie podejście miało kluczowe znaczenie dla rozwoju głębokiego uczenia się.
- Rekurencyjna sieć neuronowa (RNN) to funkcja, która nie tylko przetwarza dane wejściowe, ale także wcześniejsze dane wejściowe w czasie.
- Splotowa sieć neuronowa (CNN) analizuje dane sekcja po sekcji (czyli według splotów). Model ten jest przeznaczony do złożonych aplikacji, takich jak rozpoznawanie obrazu.
- Generacyjna sieć kontrykcyjna lub GAN to miejsce, w którym dwie sieci neuronowe konkurują ze sobą w ścisłej pętli sprzężenia zwrotnego. Rezultatem jest często stworzenie nowego obiektu.

- Wyjaśnialność opisuje techniki przejrzystości ze złożonymi modelami uczenia głębokiego.