

## Głęboka nauka

### Wprowadzenie

Głębokie uczenie się składa się z wielu hierarchicznych warstw przetwarzających informacje w sposób nieliniowy, przy czym pewna koncepcja niższego poziomu pomaga zdefiniować koncepcje wyższego poziomu. Głębokie uczenie się definiuje się jako:

Głębokie uczenie się jest klasą technik uczenia maszynowego, które wykorzystują wiele warstw nieliniowego przetwarzania informacji do nadzorowanego lub nienadzorowanego wydobywania i przekształcania cech oraz do analizy i klasyfikacji wzorców.

Płytke sztuczne sieci neuronowe nie są w stanie obsłużyć dużej ilości złożonych danych, co jest oczywiste w wielu rutynowych aplikacjach, takich jak naturalna mowa, obrazy, wyszukiwanie informacji i inne aplikacje przetwarzające informacje podobne do człowieka. Głębokie uczenie się jest zalecane dla takich aplikacji. Dzięki głębokiemu uczeniu się możliwe jest rozpoznawanie, klasyfikowanie i kategoryzowanie wzorców danych dla maszyny przy stosunkowo mniejszym wysiłku. Google jest pionierem w eksperymentowaniu z głębokim uczeniem się, które zainicjował Andrew Ng. Głębokie uczenie się oferuje wielowarstwowe przetwarzanie podobne do ludzkiego w porównaniu z płytką architekturą. Podstawową ideą głębokiego uczenia się jest zastosowanie przetwarzania hierarchicznego z wykorzystaniem wielu warstw architektury. Warstwy architektury są ułożone hierarchicznie. Dane wejściowe każdej warstwy są dostarczane do sąsiedniej warstwy po pewnym wstępnym przeszkoleniu. W większości przypadków takie wstępne szkolenie wybranej warstwy odbywa się w sposób nienadzorowany. Dogłębne uczenie się opiera się na rozproszonym podejściu do zarządzania dużymi zbiorami danych. Podejście to zakłada, że dane są generowane z uwzględnieniem wielu czynników, różnych czasów i różnych poziomów. Dogłębne uczenie się ułatwia uporządkowanie i przetwarzanie danych na różne warstwy w zależności od czasu (wystąpienia), jego poziomu lub charakteru. Głębokie uczenie się często wiąże się ze sztuczną siecią neuronową. Istnieją trzy kategorie architektury głębokiego uczenia się:

(i) Generatywne

(ii) Dyskryminujące

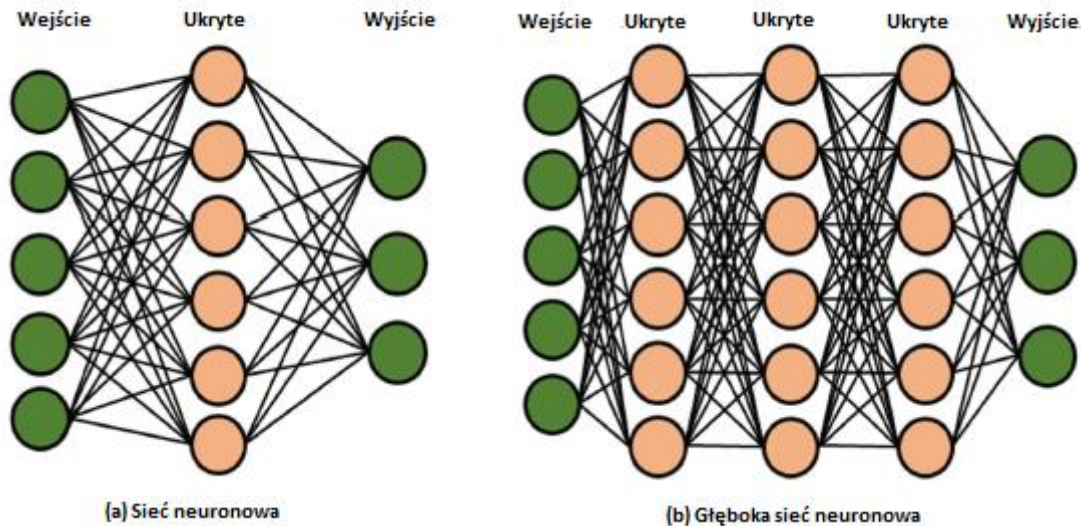
(iii) Hybrydowe architektury głębokiego uczenia się

Architektury należą do kategorii generatywnej, która koncentruje się na wstępnym szkoleniu warstwy w sposób nienadzorowany. Takie podejście eliminuje trudność szkolenia architektur niższego poziomu, które opierają się na poprzednich warstwach. Każda warstwa może być wstępnie przeszkolona, a następnie włączona do modelu w celu dalszego ogólnego strojenia i uczenia się. Takie rozwiązanie rozwiązuje problem szkolenia architektury sieci neuronowej z wieloma warstwami i umożliwia głębokie uczenie się. Architektura sieci neuronowej może mieć zdolność przetwarzania dyskryminacyjnego poprzez łączenie danych wyjściowych każdej warstwy z oryginalnymi danymi lub różnymi kombinacjami informacji, tworząc w ten sposób architekturę głębokiego uczenia. Model opisowy często traktuje wyjścia sieci neuronowej jako warunkowe rozmieszczenie we wszystkich możliwych sekwencjach znakowania dla danej sekwencji wejściowej, które będą dalej optymalizowane poprzez funkcję celu. Architektura hybrydowa łączy właściwości architektury generatywnej i dyskryminacyjnej. Typowe głębokie uczenie się można przeprowadzić w następujący sposób.

\* Zbuduj sieć składającą się z warstwy wejściowej i warstwy ukrytej z niezbędnymi węzłami

\* Trenuj sieć

- \* Dodaj kolejną ukrytą warstwę na górze wcześniej nauczonej sieci, aby wygenerować nową sieć
- \* Ponownie wyszkol sieć
- \* Powtórz dodawanie kolejnych warstw i po każdym dodaniu ponownie przeszkol sieć



### Analiza dużych zbiorów danych

Analiza dużych zbiorów danych jest wymagana do wydajnego zarządzania ogromnymi ilościami danych. Główne aspekty, które można wziąć pod uwagę przy przetwarzaniu dużych zbiorów danych, to optymalizacja na dużą skalę, wielowymiarowa obsługa danych i dynamiczna obsługa danych. Optymalizacja dotyczy znalezienia najskuteczniejszego rozwiązania problemów przy użyciu dobrze zdefiniowanych procedur i modeli. Wszyscy zajmują się problemem optymalizacji w sposób bezpośredni (systematyczny) lub pośredni (nieformalny). Problemy wymagające optymalizacji obejmują problem sprzedawcy podróży, wybór kursu z dostępnych kursów pod danym strumieniem (powiedzmy strumień naukowy) i poziom (licencjat), działania e-commerce (wybór najlepszego telefonu komórkowego za pośrednictwem różnych witryn zakupów online) itp. Optymalizacja pomaga w znalezieniu opłacalnych alternatyw dla wykonania zadania. Zwykle uważa się to za maksymalizację lub minimalizację funkcji zasobów. Niektóre przykłady to maksymalizacja zysku, minimalizacja kosztów i błędów. W przypadku domen o skończonym wymiarze dostępne są różne modele tego problemu; jednak w przypadku dużych zbiorów danych zadanie optymalizacji staje się trudnym zadaniem. W przypadku dużych zbiorów danych nie tylko wielkość transakcji jest obszerna, ale liczba zmiennych i liczba ograniczeń są również wysokie. Wręcz przeciwnie, czasami dane i ograniczenia są umiarkowane, ale ich struktura jest złożona, w taki sposób, że nie można sobie z nimi poradzić przy użyciu obecnych metod. Na przykład uczenie się funkcji z dużego repozytorium obrazów medycznych w optymalny sposób będzie trudne przy wsparciu tradycyjnej metody. Ponadto wymagałoby to ręcznego strojenia niektórych parametrów. Oprócz tradycyjnych podejść popularne stają się również metody uczenia maszynowego i optymalizacji równoległej. Wraz ze wzrostem wymiarów problemu wzrasta również złożoność. Ograniczony wymiar problemu sprawia, że problem jest łatwy do rozwiązania; jednak rozwiązanie nie jest potężne i nie zapewnia żadnej wiedzy na wysokim poziomie. Zwiększona liczba wymiarów powoduje ogromny wzrost danych, które są trudne w obsłudze, wizualizacji i rozwiązaniu. Mówi się, że z powodu wykładniczego wzrostu liczby możliwych wartości z każdym wymiarem, pełne wyliczenie wszystkich podprzestrzeni staje się trudne do uzyskania wraz ze zwiększoną wymiarowością, co jest znane jako przekleństwo wymiarowości. Dogłębne uczenie się będzie pomocne w zarządzaniu takimi wielowymiarowymi danymi oraz pomoże w grupowaniu, przetwarzaniu i

wizualizacji takich danych. Bioinformatyka, sieć czujników, rozpoznawanie wzroku i mowy to dziedziny, w których można znaleźć tak wysoko wymiarowe dane. Oprócz wielkości i struktury czas jest kolejnym ważnym czynnikiem, który zwiększa złożoność danych, a tym samym utrudnia zarządzanie danymi. Dane dynamiczne różnią się pod względem wielkości, objętości i podstawowej struktury. Wielkoskalowe i dynamiczne dane są generowane i przetwarzane w wielu obszarach, takich jak dynamika płynów, materiałoznawstwo, dynamika modułowa i systemy inspirowane biologią. Przykładową domeną jest generowanie mowy ludzkiej. Generowanie mowy ludzkiej ma strukturę hierarchiczną. Dogłębne uczenie się byłoby przydatne w modelowaniu mowy strukturalnej.

## **Różne modele głębokiego uczenia się**

### Autoencodery

Autoencoder to sztuczna sieć neuronowa zdolna do uczenia się różnych wzorców kodowania. Prosta forma autokodera przypomina wielowarstwowy perceptron zawierający warstwę wejściową, jedną lub więcej warstw ukrytych i warstwę wyjściową. Główną różnicą pomiędzy typowym perceptronem wielowarstwowym a sprzężoną siecią neuronową i autokoderem jest liczba węzłów w warstwie wyjściowej. W przypadku autokodera warstwa wyjściowa zawiera taką samą liczbę węzłów jak w warstwie wejściowej. Zamiast przewidywać wartości docelowe na podstawie wektora wyjściowego, autoencoder musi przewidywać własne dane wejściowe. Ogólny zarys mechanizmu uczenia się jest następujący.

Dla każdego wejścia  $x$

- Wykonaj przekazanie do przodu, aby obliczyć funkcje aktywacyjne dostępne na wszystkich ukrytych warstwach i warstwach wyjściowych
- Znajdź odchylenie między wartościami obliczonymi za pomocą danych wejściowych, używając odpowiedniej funkcji błędu
- Zgłoś błąd z powrotem, aby zaktualizować wagi

Powtarzaj zadanie do zadowalającego wyniku.

Jeśli liczba węzłów w ukrytych warstwach jest mniejsza niż węzłów wejściowych / wyjściowych, wówczas aktywacje ostatniej ukrytej warstwy są uważane za skompresowaną reprezentację danych wejściowych. Gdy węzły warstwy ukrytej są więcej niż warstwa wejściowa, autoencoder może potencjalnie nauczyć się funkcji tożsamości i stać się bezużyteczny w większości przypadków.

### **Deep Belief Net**

Sieć głębokiej wiary jest rozwiązaniem problemu obsługi funkcji niewypukłych i lokalnych minimów przy użyciu typowego perceptronu wielowarstwowego. Jest to alternatywny rodzaj głębokiego uczenia się składający się z wielu warstw ukrytych zmiennych połączonych między warstwami. Sieć głębokiego przekonania można postrzegać jako ograniczone maszyny Boltzmanna (RBM), w których ukryta warstwa każdej podsieci działa jako widoczna warstwa wejściowa dla sąsiedniej warstwy sieci. To sprawia, że najniższa widoczna warstwa jest zestawem szkoleniowym dla sąsiedniej warstwy sieci. W ten sposób każda warstwa sieci jest szkolona niezależnie i zachłannie. Ukryte zmienne są używane jako zmienne obserwowane do trenowania każdej warstwy głębokiej struktury. Algorytm treningowy dla sieci o głębokim przekonaniu zapewnia się następująco:

- \* Rozważ wektor wejść.
- \* Wytrenuj ograniczoną maszynę Boltzmann za pomocą wektora wejściowego i uzyskaj macierz ciężaru
- \* Trenuj dwie dolne warstwy sieci za pomocą tej macierzy wag
- \* Wygeneruj nowy wektor wejściowy za pomocą sieci (RBM) poprzez próbkowanie lub średnią aktywację ukrytych jednostek
- \* Powtarzaj procedurę aż do osiągnięcia dwóch górnych warstw sieci

Dostrajanie sieci głębokiej wiary jest bardzo podobne do perceptronu wielowarstwowego. Tak głęboko zakorzenione sieci są przydatne w modelowaniu akustycznym.

### **Konwolucyjne sieci neuronowe**

Konwolucyjna sieć neuronowa (CNN) jest kolejnym wariantem sprzężonego w przód wielowarstwowego perceptronu. Jest to rodzaj sprzężonej sieci neuronowej, w której poszczególne neurony są ułożone w taki sposób, że reagują na zachodzące na siebie obszary w polu widzenia. Głębokie sieci CNN działają poprzez modelowanie kolejno małych informacji i łączenie ich głębiej w sieci. Jednym ze sposobów ich zrozumienia jest to, że pierwsza warstwa spróbuje wykryć krawędzie i utworzyć szablony do wykrywania krawędzi. Następnie kolejne warstwy spróbują połączyć je w prostsze kształty i ostatecznie w szablony o różnych pozycjach obiektu, oświetleniu, skalach itp. Ostateczne warstwy będą pasować do obrazu wejściowego ze wszystkimi szablonami, a ostateczna prognoza jest jak ważona suma wszystkich im. Tak więc głębokie sieci CNN mogą modelować złożone warianty i zachowania, zapewniając bardzo dokładne prognozy. Taka sieć podąża za wizualnym mechanizmem żywych organizmów. Komórki w korze wzrokowej są wrażliwe na małe podregiony pola widzenia, zwane polem recepcyjnym. Podregiony są ułożone tak, aby obejmowały cały obraz pola, a komórki działają jako filtry lokalne w przestrzeni wejściowej. Algorytm wstecznej propagacji służy do trenowania parametrów każdego jądra splotu. Ponadto każde jądro jest replikowane na całym obrazie z tymi samymi parametrami. Istnieją operatory splotowe, które wydobywają unikalne cechy danych wejściowych. Oprócz warstwy splotowej sieć zawiera rektyfikowaną warstwę jednostki liniowej, warstwy pulujące w celu obliczenia maksymalnej lub średniej wartości cechy w obszarze obrazu oraz warstwę strat składającą się z funkcji strat specyficznych dla aplikacji. Rozpoznawanie obrazu i analiza wideo oraz przetwarzanie języka naturalnego są głównymi zastosowaniami takiej sieci neuronowej. W obszarze widzenia komputerowego obserwowano częste postępy w ciągu ostatnich kilku lat. Jednym z najczęściej stwierdzanych osiągnięć jest CNN. Teraz głębokie CNN stanowią rdzeń najbardziej wyrafinowanej fantazyjnej aplikacji do wizji komputerowej, takiej jak samochody samobieżne, automatyczne tagowanie znajomych na zdjęciach na Facebooku, rozpoznawanie gestów, zabezpieczenia twarzy i automatyczne rozpoznawanie tablic rejestracyjnych

### **Nawracające sieci neuronowe**

Model splotowy działa na stałej liczbie danych wejściowych, generuje wektor o stałej wielkości jako dane wyjściowe ze wstępnie określoną liczbą kroków. Sieci rekurencyjne pozwalają nam operować sekwencjami wektorów na wejściu i wyjściu. W przypadku powtarzającej się sieci neuronowej połączenie między jednostkami tworzy ukierunkowany cykl. W przeciwieństwie do tradycyjnej sieci neuronowej, rekurencyjne wejście sieci neuronowej i dane wyjściowe nie są niezależne, ale powiązane. Ponadto, rekurencyjna sieć neuronowa ma wspólne parametry na każdej warstwie. Można trenować

sieć rekurencyjną w sposób podobny do tradycyjnej sieci neuronowej, stosując metodę propagacji wstecznej. Tutaj obliczanie gradientu nie zależy od bieżącego kroku, ale również od poprzednich kroków. Wariant zwany dwukierunkową rekurencyjną siecią neuronową jest również wykorzystywany w wielu aplikacjach. Dwukierunkowa sieć neuronowa uwzględnia nie tylko poprzednie, ale także oczekiwane przyszłe wyniki. W dwukierunkowych i prostych rekurencyjnych sieciach neuronowych głębokie uczenie się można osiągnąć poprzez wprowadzenie wielu ukrytych warstw. Takie głębokie sieci zapewniają wyższą zdolność uczenia się z dużą ilością danych do nauki. Przetwarzanie mowy, przetwarzania obrazu i języka naturalnego to tylko niektóre z potencjalnych obszarów, w których można zastosować rekurencyjne sieci neuronowe.

### **Uczenie się przez wzmocnianie sieci neuronowych**

Uczenie się przez wzmocnienie jest rodzajem hybrydyzacji programowania dynamicznego i nauki nadzorowanej. Typowe elementy tego podejścia to środowisko, agent, działania, polityka i funkcje kosztów. Agent działa jako kontroler systemu; polityka określa działania, które należy podjąć; a funkcja nagrody określa ogólny cel problemu uczenia się przez wzmocnienie. Agent, który otrzymuje maksymalną możliwą nagrodę, można uznać za wykonanie najlepszej akcji dla danego stanu. Agent tutaj odnosi się do streszczenia jednostki, która może być dowolnym przedmiotem lub przedmiotem wykonującym czynności: autonomiczne samochody, roboty, ludzie, boty czatu obsługi klienta itp. Stan agenta odnosi się do pozycji agenta i jego stanu w jego abstrakcyjnym środowisku; na przykład pewna pozycja w świecie rzeczywistości wirtualnej, budynek, szachownica lub pozycja i prędkość na torze wyścigowym. Głębokie uczenie się wzmocniające stanowi obietnicę bardzo ogólnej procedury uczenia się, która może nauczyć się przydatnych zachowań przy bardzo niewielkiej ilości informacji zwrotnych. Jest to ekscytujący, ale jednocześnie wymagający obszar, który z pewnością będzie ważną częścią przyszłego krajobrazu AI.

### **Zastosowania głębokiego uczenia się w biznesie**

Analityka już od dłuższego czasu zmienia wyniki firm. Teraz, gdy coraz więcej firm opanowuje wykorzystanie analizy, zagłębia się w swoje dane, aby zwiększyć wydajność, uzyskać większą przewagę konkurencyjną i jeszcze bardziej zwiększyć swoje wyniki finansowe. Właśnie dlatego firmy chcą wdrożyć uczenie maszynowe i sztuczną inteligencję; chcą bardziej kompleksowej strategii analitycznej, aby osiągnąć te cele biznesowe. Pierwszym krokiem jest nauka, jak włączyć nowoczesne techniki uczenia maszynowego do infrastruktury danych. W tym celu wielu szuka firm, które już pomyślnie rozpoczęły proces wdrażania. Dokładnie, firmy w przestrzeni zaangażowania klienta wykorzystują sztuczną inteligencję i uczenie maszynowe do analizowania rozmów, zarówno tych, które kończą się sprzedażą, jak i tych, które tego nie robią, oraz do automatycznej identyfikacji języka, który naturalnie prowadzi do sprzedaży lub który przewiduje, kiedy sprzedaż nastąpi. Ważnym problemem jest to, czy wykorzystać cały dostępny korpus danych big data dostępny podczas analizy danych za pomocą algorytmów głębokiego uczenia. Ogólnym celem jest zastosowanie algorytmów uczenia głębokiego do szkolenia wzorców reprezentacji danych wysokiego poziomu w oparciu o część dostępnego korpusu wejściowego, a następnie wykorzystanie pozostałego korpusu wejściowego z wyuczonymi wzorcami do wyodrębnienia abstrakcji danych i reprezentacji. W kontekście tego problemu należy zbadać, jaka ilość danych wejściowych jest na ogół niezbędna do szkolenia przydatnych (dobrych) reprezentacji danych za pomocą algorytmów głębokiego uczenia, które można następnie uogólnić dla nowych danych w konkretnej domenie aplikacji dużych zbiorów danych. Jedną z najczęściej omawianych obecnie aplikacji biznesowych do głębokiego uczenia się są samochody samojezdne - koncepcja, którą stosuje każdy duży gracz, od Volkswagena po Google. Systemy te wykorzystują czujniki i sieć neuronową do przetwarzania ogromnej ilości danych. Samochód uczy się rozpoznawać przeszkody i odpowiednio reagować, zwiększając swoją wiedzę poprzez użytkowanie wykraczające poza

programowanie fabryczne. Samojezdne systemy samochodowe wykorzystują czujniki i sieć neuronową do przetwarzania ogromnej ilości danych. W końcu, mając wystarczającą ilość danych, maszyny uczą się, jak prowadzić lepiej niż ludzie. Innym powszechnym zastosowaniem jest wykrywanie obrazu i klasyfikacja obiektów, jak widać dziś na Facebooku. Firma ma więcej niż wystarczającą ilość danych na temat obrazów do pracy, dzięki czemu głębokie uczenie się w zakresie wykrywania obrazów jest bardzo dostępne. Obecnie Facebook może klasyfikować różne obiekty na obrazie z bardzo dużą dokładnością.

### **Przykład zastosowania biznesowego: Głębokie uczenie się w handlu elektronicznym**

Ogromne ilości danych są dostępne za pośrednictwem mobilnego handlu elektronicznego. Dane te umożliwiły algorytmom głębokiego uczenia się śledzenie podróży nabywcy, dzięki czemu maszyny mają teraz jasny obraz tego, jakiego rodzaju informacji o produktach szukają kupujący, kiedy podejmują decyzje zakupowe dotyczące różnych rzeczy. Ponieważ maszyny mogą jednocześnie przechowywać i odwoływać się do olbrzymich ilości informacji, stosując wyuczoną wiedzę, są w stanie przewidzieć, jakiego rodzaju zakupy konsumenci mogą dokonać przed podjęciem decyzji. Oznacza to, że śledząc aktywność użytkownika, Twoja witryna e-commerce może zaoferować kupującym opcje zakupu, które przemówią do niego na podstawie specyfikacji innych produktów, które przeglądali, ile czasu spędzili na oglądaniu każdego produktu i co zrobili przed i po obejrzeniu każdego elementu. Uwzględnia nawet porę dnia i lokalizację widza, bez względu na to, czy jest to mężczyzna, czy kobieta, oraz wszelkie inne istotne konteksty dostępne dla urządzenia. Powiedzmy, że Rita odwiedza sklep internetowy i patrzy na damskie buty do biegania Blue. Buty, na które patrzy, są ładne, ale niezupełnie tego, czego chce. Chce buta, który ma większe wsparcie łuku. Na szczęście dla Rity strona korzysta z algorytmów głębokiego uczenia się. Rita może kliknąć na zdjęcie buta, aby przejrzeć porównywalne wizualnie produkty. Produkty zalecane dla niej zostały wybrane na podstawie analizy obrazu, który kliknęła Rita, oraz jej innej aktywności, odkąd weszła na stronę. Algorytm zasugerował inne pary butów o podobnej charakterystyce koloru i kształtu. System współpracował z systemem wejścia-wyjścia, w którym działania i preferencje Rity działały jako dane wejściowe, a buty, które są jej zalecane, działają jako dane wyjściowe. Kolekcja pikseli na obrazie, które kliknęła Rita, służy jako dane wejściowe, zmuszając system do próby znalezienia podobnych zestawów danych. Oczywiście jest, że jest to skuteczny sposób nie tylko promowania swoich produktów, ale także oferowania obsługi klienta, nawet bez obecności. Klient łatwo znajduje to, czego chce, ma do dyspozycji wiele opcji, a wszystkie potrzebne informacje są łatwo dostępne. Zasadniczo Twoja witryna e-commerce staje się wirtualnym asystentem sklepu. Głębokie uczenie się oznacza więcej możliwości zwiększenia współczynników konwersji biznesowych i poprawy wizerunku marki dzięki pozytywnemu doświadczeniu klienta. Twój klient nie tylko otrzymują ukierunkowane spersonalizowane treści związane z zakupami, których dokonują, ale także oferowane im powiązane produkty, którymi mogą być zainteresowani, w oparciu o ich indywidualne zachowanie. Podczas gdy głębokie uczenie się zostało skutecznie zastosowane zadania wnioskowania wzorców, rzeczywisty cel tego pola wykracza daleko poza aplikacje specyficzne dla zadania. Ten zakres może powodować, że porównywanie różnych metod staje się coraz bardziej złożone i prawdopodobnie będzie wymagać wspólnego wysiłku ze strony społeczności badawczej. Ponadto, pomimo nieograniczonych możliwości oferowanych przez technologie głębokiego uczenia się, niektóre zadania specyficzne dla danej dziedziny mogą nie zostać bezpośrednio wzmocnione przez takie programy. Platformy dogłębnego uczenia się mogą również korzystać z funkcji inżynierskich, ucząc się bardziej złożonych reprezentacji, których zwykle brakuje w projektowanych systemach. Pomimo licznych otwartych pytań badawczych i faktu, że dziedzina ta jest wciąż w powijakach, postępy w zakresie opracowywania systemów głębokiego uczenia się będą kształtować przyszłość systemów uczenia maszynowego i systemów sztucznej inteligencji w ogóle.