

Sztuczna inteligencja (AI) nie jest nową technologią. Przez dziesięciolecia informatycy próbowali różnych podejść, aby dotrzeć do świętego Graala informatyki: inteligentnych maszyn. Chociaż wciąż jesteśmy daleko od replikowania cudów ludzkiego mózgu, aplikacje AI zaczęły wypełniać nasze codzienne życie i zasilać nasze urządzenia elektroniczne, od smartfonów po domowe systemy alarmowe. Skąd ta pozornie nagła eksplozja? Odpowiemy na to pytanie, ucząc Cię o nowoczesnej sztucznej inteligencji - w tym o jej podstawowych zasadach oraz o tym, jak i dlaczego dotarliśmy do miejsca, w którym teraz jesteśmy.

Droga do nowoczesnej sztucznej inteligencji

Jako ludzie zawsze staraliśmy się znaleźć sposoby na zrozumienie otaczającego nas świata i nagięcie natury, aby osiągnąć nasze cele. Aby to zrobić, zawsze polegaliśmy na zewnętrznych narzędziach, które wzmacniają możliwości naszego mózgu. Liczyło było prawdopodobnie pierwszym takim narzędziem, wynalezionym około 5000 do 6000 lat temu, aby pomóc ludziom w obliczeniach. Chociaż nadal jest używany w szkołach, aby pomóc dzieciom wizualizować proste operacje matematyczne, tak naprawdę nie oszczędza nas przed trudem ich rzeczywistego wykonywania. Na pierwsze maszyny, które potrafiły automatycznie dodawać i odejmować liczby, musieliśmy czekać do lat 60. XX wieku. Od tego czasu komputery przeszły długą drogę, ale w głębi ich możliwości wciąż są dość proste: wykonują obliczenia dokładnie tak, jak polecił im (ekspert) człowiek. Mało jest w nich „inteligencji”. Dwa słowa sztuczna i inteligencja zostały po raz pierwszy połączone 31 sierpnia 1955 roku, kiedy profesor John McCarthy z Dartmouth College wraz z M.L Minsky z Harvard University, N. Rochesterem z IBM i C. E. Shannonem z Bell Telephone Laboratories zapytali Fundację Rockefellera o sfinansowanie lata badań nad sztuczną inteligencją. Ich propozycja stwierdzała, co następuje:

„Proponujemy, aby latem 1956 roku w Dartmouth College w Hanover, New Hampshire, przeprowadzono dwumiesięczne, 10-osobowe badanie sztucznej inteligencji … . Podjęta zostanie próba znalezienia sposobu na zmuszenie maszyn do używania języka, tworzenia abstrakcji i pojęć, rozwiązywania rodzajów problemów zarezerwowanych obecnie dla ludzi i samodoskonalenia. Uważamy, że w jednym lub kilku z tych problemów można dokonać znacznego postępu, jeśli starannie dobrana grupa naukowców będzie nad nim pracować razem przez lato. „

Naukowcy wiedzieli, że zajmowanie się inteligencją jako całością było zbyt trudnym wyzwaniem, zarówno ze względu na ograniczenia techniczne, jak i nieodłączną złożoność zadania. Zamiast rozwiązywać szerokie pojęcie inteligencji, postanowili skupić się na podproblemach, takich jak język. Później aplikacje te nazwano by wąską sztuczną inteligencją. Sztuczna inteligencja zdolna do dopasowania lub przewyższenia ludzkich możliwości zostałaaby zamiast tego nazwana ogólną sztuczną inteligencją. Innymi słowy:

* Ogólna sztuczna inteligencja (lub silna sztuczna inteligencja) - program sztucznej inteligencji zdolny do radzenia sobie z każdym rodzajem przedstawionego zadania. Jest to podobne do niezwykle zaradnego człowieka i możesz myśleć o nim jako o robocie z Terminatora (lub, miejmy nadzieję, o bardziej spokojnej wersji tego).

* Wąska sztuczna inteligencja - program sztucznej inteligencji zdolny do rozwiązania pojedynczego, dobrze zdefiniowanego zadania. Może być szeroki (rozpoznawanie obiektów ze zdjęć) lub bardzo szczegółowy (przewidywanie, którzy klienci, którzy kupili produkt A, z większym prawdopodobieństwem kupią również produkt B). Oznacza to jedno zadanie na raz, a nie żadne inne: sztuczna inteligencja, która rozpoznaje koty na obrazach, nie może tłumaczyć z angielskiego na włoski i odwrotnie.

Ogólna sztuczna inteligencja wciąż jest daleko, badacze wciąż nie wiedzą, kiedy w końcu ją zdobędziemy. Niektórzy twierdzą, że nigdy tam nie dotrzemy. Mimo że ogólna sztuczna inteligencja jest nadal odległym, rozmytym snem, wielu ludzi ma to na myśli, gdy w wiadomościach pojawia się sztuczna inteligencja. Jeśli byłeś jedną z tych osób i jesteś teraz rozczarowany, że ogólnej sztucznej inteligencji jeszcze nie ma, nie rozpaczaj. Wąskie aplikacje AI nadal są w stanie tworzyć ogromną wartość. Na przykład sztuczna inteligencja, która może wykryć raka płuc, jest wąskim zastosowaniem, ale mimo to niezwykle przydatna. Wyniki badań latem 1956 r. w Dartmouth były tak interesujące, że wywołały wśród uczestników falę podekscytowania i nadziei. Entuzjazm naukowców rozprzestrzenił się na rząd USA, który zaczął intensywnie finansować badania nad konkretnym zastosowaniem: tłumaczeniem na język angielski/rosyjski. Znalezienie godnych zaufania rosyjskich tłumaczy nie było łatwe w środku zimnej wojny. Po pierwszych kilku latach pracy komitet rządowy opracował niesławny raport Komitetu Doradczego ds. Automatycznego Przetwarzania Języka (ALPAC) z 1966 roku. W dokumencie znalazły się opinie wielu badaczy na temat stanu badań nad sztuczną inteligencją. Większość nie była zbyt pozytywna:

„Wczesne tłumaczenia maszynowe prostego lub wybranego tekstu … były tak samo zwodniczo zachęcające, jak „tłumaczenia maszynowe” tekstów ogólnonaukowych były jednolicie zniechęcające … Nikt nie może oczywiście zagwarantować, że nie osiągniemy nagle lub przynajmniej szybko tłumaczenia maszynowego, ale uważamy, że jest to bardzo mało prawdopodobne.

... nie ma natychmiastowej ani przewidywalnej perspektywy przydatnego tłumaczenia maszynowego.,,

Raport ALPAC wyznacza początek okresu zwanego pierwszą zimą AI: wstrzymano publiczne finansowanie badań nad sztuczną inteligencją, podekscytowanie opadło, a naukowcy skupili się na innych dziedzinach. Zainteresowanie sztuczną inteligencją osłabło do lat 80., kiedy prywatne firmy, takie jak IBM i Xerox, zaczęły inwestować w nową wiosnę AI. Nowe nadzieje podsycała technologia zwana systemami eksperckimi: programy komputerowe, które kodują wiedzę człowieka-eksperta w określonej dziedzinie w formie precyzyjnych, reguł jeśli to. Przykład pomoże Ci zrozumieć, w jaki sposób zaprojektowano systemy eksperckie. Załóżmy, że chcesz zbudować system sztucznej inteligencji, który mógłby zastąpić gastroenterologa. Tak się to robi z systemem eksperckim: prosisz lekarza, aby z niezwykłą precyzją opisał, w jaki sposób podejmuje decyzje dotyczące pacjentów. Następnie prosisz programistę o drobiazgową transformację wiedzy lekarza i diagnozy które przepływają do reguł „jeśli-to”, które mogą być zrozumiane i wykonywane przez komputer. Niezwykle uproszczona wersja wyglądałaby mniej więcej tak:

Jeśli pacjent ma bóle brzucha i temperatura ciała jest wysoka,
wtedy pacjent ma grypę.

Jeśli pacjent ma ból brzucha i zjadł przeterminowane jedzenie
wtedy pacjent ma zatrucie pokarmowe.

I tak dalej. Gdy wiedza lekarza zostanie zakodowana w oprogramowaniu i pojawi się pacjent, oprogramowanie podąża tą samą ścieżką decyzji co lekarz i (miejmy nadzieję) stawia tę samą diagnozę. Takie podejście ma kilka problemów:

* Słaba adaptacyjność - jedynym sposobem na ulepszenie oprogramowania jest powrót do deski kreślarskiej z informatykiem i ekspertem (w tym przypadku z lekarzem).

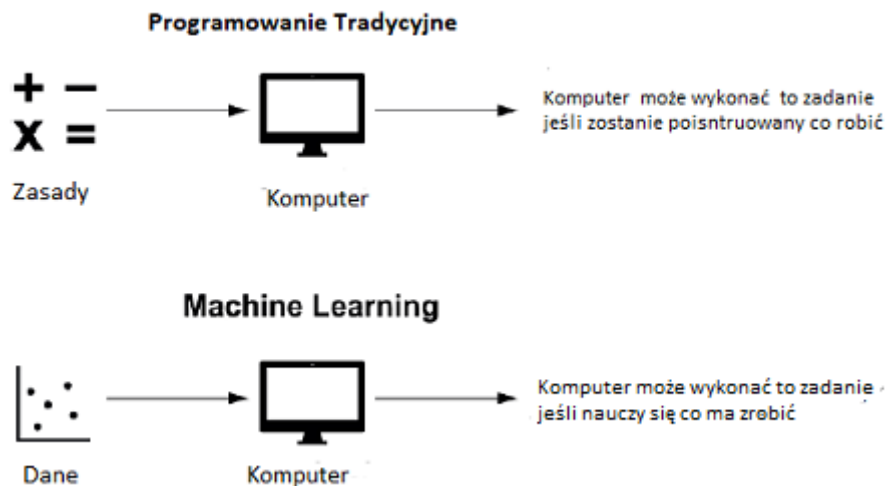
* Ekstremalna kruchość - system zawiedzie w sytuacjach, które nie były częścią oryginalnego projektu. Co się stanie, jeśli pacjenta boli brzuch, ale temperatura ciała jest normalna i nie jadł zepsutego jedzenia?

* Trudne w utrzymaniu - Złożoność takiego systemu jest ogromna. Kiedy zestawia się tysiące reguł, ulepszanie ich lub zmienianie jest niezwykle skomplikowane, powolne i kosztowne. Czy kiedykolwiek pracowałeś z ogromnym arkuszem programu Microsoft Excel i próbowałeś znaleźć podstawową przyczynę błędu? Wyobraź sobie arkusz Excela 100 razy większy.

Systemy eksperckie okazały się komercyjną porażką. Pod koniec lat 80. wiele firm, które je rozwijały, wypadło z rynku, rozpoczynając drugą zimę AI. Dopiero na początku XXI wieku pojawiła się następna generacja sukcesów sztucznej inteligencji, napędzana starym pomysłem, który znów stał się nowy: uczenie maszynowe.

Silnik rewolucji AI: uczenie maszynowe

Pierwsza definicja uczenia maszynowego pochodzi z 1959 r., autorstwa amerykańskiego pioniera sztucznej inteligencji Arthura Samuela: *Uczenie maszynowe to dziedzina nauki, która daje komputerom możliwość uczenia się bez wyraźnego programowania. Kluczowymi elementami są tutaj uczenie się i bez wyraźnego programowania. Skupmy się najpierw na tym drugim. Jawne programowanie komputera oznacza zdefiniowanie zasad i instrukcji, którymi musi się kierować, aby wykonać określone zadanie. To właśnie robią inżynierowie oprogramowania, kiedy piszą oprogramowanie, które obsługuje Twoje codzienne zadania, takie jak rozliczanie podatków lub wypełnianie arkuszy kalkulacyjnych. Osoby bez doświadczenia w programowaniu często czują się, jakby inżynierowie oprogramowania byli potężnymi stworzeniami, które potrafią naginać maszyny do swojej woli. Niestety nie zawsze jest tak łatwo. Spróbuj pomyśleć o różnych decyzjach, które podejmujesz, wykonując kilka trywialnych czynności: czy możesz wyjaśnić, w jaki sposób rozpoznajesz znajomych, gdy ich widzisz? Wszystkie decyzje podejmowane w ułamku sekundy podczas jazdy? Czy możesz wymienić wszystkie zasady gramatyki angielskiej, które stosujesz podczas rozmowy? Jeśli nie potrafisz dokładnie wyjaśnić, jak coś robisz, nie ma szans, abyś zlecił to komputerowi. Samuel zaproponował zastąpienie „instruowania komputerów” „dawaniem im możliwości uczenia się”. Jeśli się nad tym zastanowisz, uczenie się zamiast wykonywania instrukcji jest (przypadkowo?) tym, co ludzie robią cały czas. Nasze matki i ojcowie nie uczą nas swojego języka ojczystego, podając gramatykę w młodym wieku jednego roku. Po prostu mówią do nas naturalnie, a my uczymy się z ich przykładu, stosując tysiące reguł gramatycznych, nawet o tym nie wiedząc. W rzeczywistości nasz mózg jest w stanie automatycznie wydobywać reguły, zanim będzie w stanie racjonalnie rozumieć gramatykę w szkole! Nawet nam, ludziom, wygląda na to, że uczenie się reguł na przykładach może być łatwiejsze niż opowiadanie o nich. W ten sam sposób, w jaki uczymy się z doświadczenia, techniki uczenia maszynowego (ML) umożliwiają komputerom uczenie się na podstawie danych. Zróbmy to bardziej konkretnie na przykładzie klasycznej zabawki: nauczenie komputera odróżniania psów od kotów na zdjęciach. Gdybyś musiał nauczyć dziecko wykonywania tego zadania, nie wzięłbyś książki weterynaryjnej i nie zacząłbyś czytać o różnicach w kształcie uszu lub kolorze futra. Zamiast tego prawdopodobnie po prostu wskażesz im kilka zdjęć i pozwolisz im mózgowi wykonać swoją magię. Rozwiązanie problemu „psa lub kota” w ML jest podobne do naszych doświadczeń związanych z nauką w dzieciństwie. Karmimy komputer tysiącami zdjęć kotów i mówimy „to są koty”, a następnie tysiącami zdjęć psów i mówimy „to są psy”. Na koniec pozwalamy mu automatycznie ustalić różnicę między dwoma zwierzętami. Nie musimy wyjaśniać kluczowych elementów, które odróżniają psy od kotów. Dobra aplikacja ML uczy się tego rozumieć na podstawie przykładów, które otrzymuje. Rysunek pokazuje różnicę między tradycyjnym programowaniem a uczeniem maszynowym.*



Możesz zacząć rozumieć, dlaczego ML nie mogło rozkwitnąć przed 2000 rokiem. Głównym składnikiem tego zestawu technik są dane, a internet znacznie ułatwił zbieranie danych. Innym kluczowym elementem ML jest moc obliczeniowa: uczenie się na podstawie danych nie odbywa się za darmo, a komputery potrzebują szybkich procesorów do wykonania tego zadania. Dzięki przetwarzaniu w chmurze i zwiększeniu mocy obliczeniowej dostęp do potężnych komputerów jeszcze nigdy nie był tak łatwy i tani. Aby pokazać, jak wiele zmieniło się w ciągu zaledwie kilku lat, zapytaliśmy Alexa Waibela, jednego z pionierów sztucznej inteligencji w rozpoznawaniu mowy i jednego z pierwszych zatrudnionych w zespole AI Facebooka, jak inaczej wyglądała praca nad ML 20 lat temu. Najpotężniejszy komputer, którego mógł używać na początku 2000 roku, był wielkości mieszkania, kosztował kilka milionów i musiał go wynająć, aby trenować swoje modele. Dziś ma na biurku znacznie większą moc obliczeniową za kilka tysięcy dolarów. Twój telefon jest prawdopodobnie potężniejszy niż to, co najlepsi badacze mieli do dyspozycji zaledwie 20 lat temu. Dostępność danych i tania moc obliczeniowa stworzyły idealne środowisko do rozkwitu uczenia maszynowego. Rzeczywiście, wiele (większość) najfajniejszych aplikacji skierowanych do konsumentów w to, co dziś nazywamy sztuczną inteligencją, w dużej mierze opiera się na ML: wirtualnym asystencie głosowym Siri, Tłumaczu Google, samojezdnym samochodach i wielu innych. Wracając do historii sztucznej inteligencji, wydaje się, że ML to silnik, który napędzał dzisiejszą eksplozję sztucznej inteligencji, w końcu przynosząc nadzieję po ostatniej zimy AI w latach 80. XX wieku. W rzeczywistości sukces współczesnej sztucznej inteligencji był tak zależny od technik ML, że ludzie często są zdezorientowani różnicą między nimi. Czym więc jest sztuczna inteligencja? Dowiedzmy się.

Czym w końcu jest sztuczna inteligencja?

W naszym życiu ostatni nieustannie spotykamy ludzi o różnych opiniach na temat definicji AI, data science i ML. Chociaż wielu jest dość upartych, niewielu jest w stanie obronić swoją pozycję. Rzeczywiście, znalezienie uniwersalnej definicji AI nie jest tak trywialne, jak mogłoby się wydawać. Idąc po jej nazwie, moglibyśmy spróbować zdefiniować sztuczną inteligencję, znajdując ludzkie cechy, które kojarzą nam się z inteligencją. Kiedy już zgodzimy się, co sprawia, że ludzie są inteligentni, możemy powiedzieć, że każdy komputer, który robi to samo, jest sztuczną inteligencją. To ma sens, prawda? Chociaż jest to powszechne podejście, rozpada się nawet przy prostych scenariuszach. Na przykład człowiek, który w ułamku sekundy potrafi podzielić 13,856 przez 13 z dokładnością do dziesiątej liczby po przecinku, z pewnością zostałby nazwany inteligentnym, ale jego sztucznym odpowiednikiem jest

kieszonkowy kalkulator za 20 PLN, którego nikt nie odważyłby się nazwać sztuczną inteligencją. Jednocześnie nigdy nie nazwalibyśmy kogoś inteligentnym tylko dlatego, że potrafi jeździć w dużym natężeniu ruchu, jednak samochód autonomiczny jest powszechnie uważany za jedną z najtrudniejszych form sztucznej inteligencji, nad którą obecnie pracuje przemysł technologiczny. Nie powinno nas dziwić, jak trudno jest zdefiniować inteligencję; w końcu filozofowie i naukowcy dyskutują o tym od wieków. Nie tylko mamy różne wagi do pomiaru inteligencji ludzi i maszyn, ale także wydaje się, że dość szybko zmieniamy zdanie na temat tego, czym jest sztuczna inteligencja, a co nie. Weźmy za przykład Paula Grahama, założyciela Y Combinator, najbardziej udanego akceleratora startupów z Doliny Krzemowej i prawdopodobnie jednego z najbardziej patrzących w przyszłość ludzi w dziedzinie technologii. W 2002 roku Graham napisał esej, proponując nowe rozwiązanie do wykrywania wiadomości spamowych. W tamtych czasach poczta e-mail dopiero zaczynała działać, a spam (niechciana poczta e-mail) był jednym z najpoważniejszych zagrożeń dla powszechnego korzystania z Internetu przez osoby nie będące technikami. Wydaje się to trudne do wyobrażenia teraz, ale najlepsi informatycy byli zajęci pisaniem złożonych reguł, aby komputery automatycznie sortowały reklamy Viagry. W swoim esej Graham pomyślał o nowym podejściu opartym na ML, które nauczyłoby się klasyfikować wiadomości e-mail poprzez przetwarzanie tysięcy „dobrych” i spamowych wiadomości e-mail. Proste oprogramowanie Paula nauczyło się rozpoznawać spam lepiej niż złożone reguły wymyślane przez inżynierów. Szybko do przodu o 20 lat, a automatyczne wykrywacze spamu to tak nudna technologia, że wylecielibyśmy z pokoju, gdybyśmy odważyli się nazwać to sztuczną inteligencją. W rzeczywistości wydaje się, że sztuczna inteligencja polega na opanowaniu zadań, które nasze wyobrażenia podpowiadają, że komputery nie powinny być w stanie. Kiedy przyzwyczajamy się do technologii w naszym codziennym życiu, usuwamy honorową odznakę AI i zaczynamy nazywać ją tylko oprogramowaniem komputerowym. Jest to dobrze zbadane zjawisko zwane efektem AI. Ze względu na efekt sztucznej inteligencji, bramki dla tego, co nazywamy sztuczną inteligencją, poruszają się tak szybko, jak poprawia się technologia. Definicja sztucznej inteligencji, którą czerpiemy z tych rozważań, to „tymczasowa etykieta dla oprogramowania, które robi coś fajnego i zaskakującego, dopóki się do tego nie przyzwyczajamy”. Nie wiemy o Tobie, ale to po prostu nie wydaje się satysfakcjonującą definicją. Mamy nadzieję, że przekonaliśmy Cię, że niezwykle trudno jest znaleźć definicję, która uszczęśliwi wszystkich i będzie aktualna wraz z rozwojem technologii. Mając na uwadze efekt sztucznej inteligencji, postanowiliśmy uniknąć wąskiej definicji sztucznej inteligencji, która nagradza „efektywne” aplikacje tylko po to, aby je porzucić, gdy szum zniknie. Przyjmujemy szerszą definicję, która obejmuje mniej krzykliwe aplikacje. Oto nasza definicja sztucznej inteligencji:

Oprogramowanie, które rozwiązuje problem bez wyraźnej instrukcji ludzkiej.

Jak widać, nasza definicja skupia się na wyniku technologii, a nie na konkretnych technikach użytych do jej zbudowania. Niektórzy się z tym nie zgodzą, ponieważ jest to prawie równoznaczne z tym, co powiedzieliśmy o uczeniu maszynowym we wcześniejszej części. Prawda jest taka, że uczenie się jest inteligentną cechą i chociaż ML jest tylko narzędziem, to jest narzędziem, które stoi za 99% udanych aplikacji, które dzisiaj nazywamy sztuczną inteligencją. Może się to zmienić w przyszłości, ale nie widzimy na horyzoncie żadnych nowych podejść, które miałyby taką samą obietnicę jak ML. Dlatego każda aplikacja AI, którą omówimy, opiera się na ML: jest to po prostu najdokładniejszy obraz dzisiejszego krajobrazu AI i najbliższej przyszłości. Mamy teraz jasny obraz tego, czym jest ML, działającą definicję nowoczesnej sztucznej inteligencji oraz pewne spojrzenie na ewolucję tych terminów. Brakuje nam trzeciego modnego hasła, o którym prawdopodobnie słyszałeś: data science. Nauka o danych (DS) to szeroka, multidyscyplinarna dziedzina, która wykorzystuje naukowe metody i procesy do analizowania danych i wydobywania spostrzeżeń. Techniki ML to tylko niektóre z narzędzi w zestawie narzędzi DS. W praktyce, kiedy ludzie odnoszą się do projektu z zakresu analizy danych, często mają na myśli coś statycznego: wydobywanie spostrzeżeń z danych i przedstawianie ich w

formie prezentacji lub raportu. Z drugiej strony sztuczna inteligencja jest częściej używana w kontekście oprogramowania na żywo. Na przykład analizowanie danych o ruchu w celu zaprojektowania nowego planu urbanistycznego dla miasta w celu zminimalizowania zatorów komunikacyjnych prawdopodobnie wchodzi w zakres nauki o danych. Jeśli jednak użyjesz tych samych danych do kontrolowania ruchu w czasie rzeczywistym i kierowania samochodów na mniej zatłoczone trasy, większość ludzi powie, że projekt dotyczy sztucznej inteligencji. W pierwszym przypadku wynikiem projektu jest raport, a w drugim jest to oprogramowanie „na żywo”, które działa 24 godziny na dobę, 7 dni w tygodniu. Należy pamiętać, że ten podział jest w większości konwencjonalny: tam naprawdę nie ma sztywnych i szybkich zasad dotyczących sztucznej inteligencji i nauki o danych. Tabela podsumowuje różnice, jakie widzimy.

Sztuczna inteligencja: Nauka o danych

Automatyzuje zadania lub przewiduje przyszłe zdarzenia na podstawie danych. : Generuje wnioski na podstawie danych.

Jest powszechnie używany „na żywo”: Stale opracowuje nowe dane i dostarcza odpowiedzi. :

Jest zwykle „jednorazowy”: Daje pewne spostrzeżenia, które wpływają na decyzje.

Zwykle ma postać oprogramowania. : Zwykle ma formę prezentacji lub raportu.

Mamy nadzieję, że te sekcje pomogły w wyjaśnieniu niektórych powszechnie błędnie rozumianych terminów i stworzyły kontekst dla tych technologii. Teraz możesz zacząć uczyć się podstawowych zasad sztucznej inteligencji, tego, co potencjalnie możesz z nią zrobić i jak wprowadzić tę transformacyjną technologię do swojej organizacji. W następnej sekcji wyjaśnimy etapy tej podróży i sposób, w jaki ten tekst poprowadzi Cię przez nie.

* Sztuczna inteligencja ma długą historię sukcesów i porażek, sięgającą lat 50. XX wieku.

* Ogólna sztuczna inteligencja to mrzonka posiadania wszechwiedzącej maszyny. Wszystkie aplikacje AI, jakie mamy dzisiaj, są raczej wąskie; skupiają się na konkretnych zadaniach.

* Uczenie maszynowe jest obecnie powszechnym sposobem wdrażania sztucznej inteligencji i opiera się na umożliwieniu maszynom samodzielnego uczenia się na podstawie danych.

* Nauka o danych jest powiązana ze sztuczną inteligencją i uczeniem maszynowym, ale koncentruje się bardziej na wydobywaniu spostrzeżeń niż na trwałej inteligencji.

Sztuczna inteligencja dla podstawowych danych biznesowych

Tu omówimy podstawowe dane biznesowe , czyli rodzaj danych, który jest najbliższy propozycji wartości Twojej firmy. Podstawowe dane biznesowe to sekretny sos, który napędza Twoją organizację. Są to dane o zamówieniach gromadzone przez operacje komercyjne, interakcje międzyludzkie śledzone przez sieci społecznościowe oraz pomiary ciśnienia krwi śledzone przez start-up medyczny. Ponieważ podstawowe dane biznesowe znajdują się w sercu organizacji, są one również głównym kandydatem do pierwszych aplikacji AI. Wykonywanie sztucznej inteligencji na podstawowych danych biznesowych jest jak oszustwo: dane są tak blisko silnika generującego wartość organizacji, że udane projekty mają prawie gwarancję, że wywrą dramatyczny wpływ. Aby zaostrzyć Twój apetyt, pokrótce wspomnijmy o dwóch studiach przypadków, które znajdziesz na końcu. Pierwsza dotyczy inicjatywy Jima Gao, byłego pracownika Google odpowiedzialnego za operacje w jego centrach danych. Przyjrzał się danym zebranych z dużych systemów klimatyzacyjnych używanych do chłodzenia gigantycznych komputerów Google i pomyślał o wykorzystaniu uczenia maszynowego do optymalizacji ich zużycia. Rezultatem było

cięcie rachunków o 40% dla giganta technologicznego. Drugi przypadek dotyczy Square, firmy świadczącej usługi płatnicze z siedzibą w San Francisco. Square przetwarzał płatności kartą kredytową dla małych firm i zdał sobie sprawę, że wszystkie gromadzone dane można wykorzystać do oferowania małym firmom dostosowanych i niskiego ryzyka pożyczek. Korzystając z modeli uczenia maszynowego, Square stworzył zupełnie nową linię biznesową i zajął się błękitnym oceanem pożyczek dla małych firm. Jakość usług nie ma sobie równych: automatyzując proces kredytowania, Square może wpłacać pożyczki na konto bankowe klienta już w jeden dzień roboczy po złożeniu wniosku. W rezultacie Square pożyczył ponad 3 miliardy dolarów w ciągu czterech lat, przy wyjątkowo niskich wskaźnikach zaległości wynoszących 4%. Pod koniec dowiesz się, w jaki sposób te firmy umożliwiły te zmiany.

Uwolnij sztuczną inteligencję na podstawowych danych biznesowych

Podstawowe dane biznesowe definiujemy jako „dane mające bezpośredni wpływ na górną lub dolną linię organizacji”. Podstawowe dane wyglądają bardzo różnie w zależności od tego, co robi Twoja organizacja: historia koszyka dla operacji e-commerce, pomiary fizyczne dla organizacji inżynierskiej i zachowanie pacjentów w firmie opieki zdrowotnej. Niezależnie od formy, dane podstawowe są cenne, ponieważ opisują zdarzenia i wzorce, które mają bezpośredni wpływ na wyniki organizacji, i łatwo jest przypisać im wartość pieniężną. Wspomnieliśmy już o dwóch studiach przypadków, które czekają na Ciebie na końcu tego rozdziału. Przyjrzyjmy się, dlaczego dane wykorzystywane przez te dwie firmy są podstawowymi danymi biznesowymi:

* Firma taka jak Google polega na ogromnych centrach danych, aby oferować swoje usługi (przetwarzanie wyszukiwań internetowych, przechowywanie zdjęć, kierowanie wiadomości e-mail itd.). Prawdopodobnie jedynym zmiennym kosztem dla Google jest energia zużywana na chłodzenie komputerów w centrach danych. Dlatego dane o instalacjach chłodniczych mają kluczowe znaczenie dla działalności Google, ponieważ są bezpośrednio powiązane z jednym z głównych kosztów.

* Podstawowym produktem Square jest rozwiązanie do obsługi punktów sprzedaży (POS). Dzięki temu produktowi Square przetwarza wszystkie płatności trafiające do jego klientów. Ponieważ misją Square jest wzmacnianie pozycji małych firm, dane transakcyjne są ściśle powiązane z tą wizją i są cenne dla klientów.

Dobrym sposobem na przyjrzenie się wartości danych jest zastanowienie się nad miernikiem, który nazwiemy gęstością danych w złotych: jak bardzo dane wpływają na górną lub dolną linię organizacji. Podstawowe dane biznesowe mają duże zagęszczenie w złotych: każde zamówienie e-commerce, zlecenie lub transakcja finansowa ma bezpośredni wpływ na wyniki lub wyniki finansowe. Gdy odchodzisz od podstawowej propozycji wartości swojej organizacji, gęstość gromadzonych danych odpowiednio się zmniejsza. Rejestrowanie wizyt w Twojej witrynie może być cenne, ale nie tak bardzo, jak śledzenie napływających zamówień. W przypadku szpitala dane z call center mogą być przydatne, ale nie tak bardzo jak dane pacjentów. Dlatego zdecydowaliśmy się rozpocząć od danych o największej gęstości ze wszystkich: podstawowych danych biznesowych. Często widzimy, że podstawowe dane biznesowe przybierają ustrukturyzowaną formę, podobnie jak uporządkowane wiersze i kolumny arkusza kalkulacyjnego Microsoft Excel. W kategoriach inżynierskich nazywamy ten rodzaj danych ustrukturyzowanymi. Inne przykłady danych strukturalnych to raporty pogodowe, pomiary procesów fizycznych, transakcje finansowe, większość rynków oraz wskaźniki łańcucha dostaw i magazynowania. Z reguły wszystko, co można załadować w programie Excel, prawdopodobnie będzie objęte parasolem danych strukturalnych. Inne rodzaje danych trudniej zmieścić w zgrabnych kolumnach Excela: pomysły o zdjęciach, nagraniach głosowych lub tekście w książce. Ważne jest, aby zrozumieć, że te same informacje mogą istnieć zarówno w formie ustrukturyzowanej, jak i nieustrukturyzowanej. Rozważmy na przykład następujący sposób rejestracji diagnozy medycznej:

Pacjent Gianluca Mauro ma ciężkie zapalenie stawu barkowego; terapia polega na przyjmowaniu dwóch tabletek kortyzonu dziennie przez pięć dni. Te same informacje można zapisać w uporządkowany sposób, jak w tabeli .

Pacjent : Diagnoza : Obszar : Leki : Częstotliwość (razy/dzień) : Długość terapii (dni)

Gianluca Mauro : Ciężkie zapalenie : Staw barkowy : Kortyzon : 2:5

Informacje są takie same, ale w pierwszym przypadku są przedstawione w sposób nieustrukturyzowany (tekst), podczas gdy w drugim są ustrukturyzowane. Dane strukturalne są znacznie łatwiejsze do przetwarzania przez komputery niż dane niestrukturalne. Późniejsze części obejmują techniki sztucznej inteligencji dla danych nieustrukturyzowanych, w tym obrazów i języka pisanego. Na razie pozostaniemy przy danych ustrukturyzowanych, wiedząc, że większość podstawowych danych biznesowych należy do tej kategorii.

Korzystanie ze sztucznej inteligencji z podstawowymi danymi biznesowymi

Teraz, gdy wiesz, jak znaleźć i rozpoznać podstawowe dane biznesowe, przyjrzyjmy się, co może z nimi zrobić sztuczna inteligencja. Ponieważ wierzymy w uczącą moc opowieści, opracowaliśmy uproszczony przykład, który dotrzyma Ci towarzystwa przez całą pierwszą część. FutureHouse to fikcyjna firma, która prowadzi internetowy rynek nieruchomości, na którym właściciele domów mogą reklamować swoje domy na sprzedaż i, miejmy nadzieję, przyciągnąć zainteresowanych kupujących. FutureHouse zawsze priorytetowo traktowała obsługę klienta i zatrudnia agentów, którzy mogą zaoferować swoje wsparcie, aby pomóc sprzedawcom ocenić cenę ich domu. Kupujący korzystają z serwisu w poszukiwaniu domu swoich marzeń. Wybraliśmy ten przykład, ponieważ będzie on znany wielu czytelnikom, a jednocześnie da nam możliwość poznania wielu aspektów krajobrazu sztucznej inteligencji. Miłym bonusem jest to, że rynek mieszkaniowy jest typowym przykładem w literaturze ML. Jeśli zdecydujesz się zagłębić w szczegóły techniczne ML i być może napisać własny kod, w sieci znajdziesz mnóstwo odniesień. Zaczniemy od wyjaśnienia, jak działają rynki mieszkaniowe, a następnie przedstawimy zastosowanie ML do tego biznesu. Użyjemy tego opisu jako pretekstu do wprowadzenia kluczowej terminologii i pojęć, a na koniec rozszerzymy te pojęcia na inne, bardziej ogólne zastosowania.

Przykład rynku nieruchomości

Agencje nieruchomości znajdują się w centrum rynków mieszkaniowych. Łączą sprzedawców i kupujących i pomagają im ustalić odpowiednią cenę za ich nieruchomość. Oto jak przebiega typowa transakcja sprzedaży domu pod uważnym okiem pośrednika w handlu nieruchomościami:

1. Klient przychodzi i chce sprzedać swój dom. Zanim wprowadzą go na rynek, chcą profesjonalnej opinii o tym, ile jest wart.
2. Pośrednik sprawdza dom i ustala cenę na podstawie jego powierzchni, wieku, wliczonych w cenę urządzeń, innych ofert w okolicy i tak dalej.
3. Dom zostaje wymieniony.
4. Potencjalni nabywcy znajdują ofertę i przychodzą na wydarzenie otwarte. Jeden szczęśliwy nabywca w końcu może kupić nieruchomość.

FutureHouse zbiera dane na każdym etapie tego procesu. Oto jego główne zasoby danych:

- * Zdjęcia wszystkich domów na sprzedaż
- * Opisy domów (metryczna powierzchnia, lokalizacja itp.)
- * Historia odwiedzin i kliknięć dla wszystkich użytkowników witryny
- * Zapisy historyczne nieruchomości sprzedanych w regionie
- * Recenzje domów pozostawionych przez osoby, które je odwiedziły
- * Transkrypcje negocjacji między agentem a kupującym
- * Artykuły prasowe o zbliżającej się budowie w regionie

Wszystkie te źródła danych nie mają tej samej wartości; ich gęstość w dolarach jest inna. Na przykład artykuły prasowe są bardzo ogólnym zbiorem danych, do którego dostęp ma każda inna firma. Dlatego, chociaż może być wartościowe, nie jest podstawą FutureHouse. Odwiedziny w witrynie, obrazy i recenzje mogą być ważne, ponieważ są specyficzne i unikalne dla firmy, a zatem ich gęstość w dolarach jest wyższa niż w przypadku artykułów prasowych. Nie mają jednak bezpośredniego wpływu na powodzenie biznesu. Historyczne zapisy dotyczące sprzedanych domów są znacznie bardziej powiązane z wynikami firmy niż wszystkie inne źródła danych, ponieważ tradycyjny model biznesowy pośredników w handlu nieruchomościami to opłata za sukces za sprzedany dom: są to podstawowe dane biznesowe. Załóżmy, że chcemy wykorzystać te dane do zbudowania opartego na sztucznej inteligencji narzędzia do oceny nieruchomości, które automatycznie przewiduje najlepszą cenę domów notowanych na platformie. Wycena jest zwykle wykonywana przez agenta, który ma doświadczenie w okolicy i podobnych domach. W przeszłości nadzorowali sprzedaż wielu nieruchomości, rozwijając głęboką intuicję, jaka byłaby uczciwa cena nieruchomości. To jest zadanie, które chcemy przejąć AI. Spróbujmy wymyślić bardziej wyczerpującą listę czynników, które agent może wziąć pod uwagę przed zasugerowaniem ceny sprzedaży:

- * Powierzchnia kwadratowa
- * Numer piętra
- * Liczba sypialni
- * Liczba łazienek
- * Rok budowy
- * Czy ma basen?
- * Czy ma garaż?
- * Efektywności energetycznej
- * Lokalizacja
- * Jakość sąsiedztwa
- * Transport publiczny

Jeśli zapytasz, jak agent wymyślił cenę, prawdopodobnie powiedzą coś w stylu „z doświadczenia”. Mogą odnosić się do standardowych wytycznych, takich jak średnia cena za metr kwadratowy w tej okolicy, ale trudno im wyrazić, w jaki sposób inne czynniki wpłynęły na ich decyzję. Jeśli nie potrafisz wyjaśnić swojego procesu myślowego, nie możesz nawet zakodować go w programie komputerowym.

Agent uczy się, jak wyceniać domy, patrząc na ceny sprzedaży wielu nieruchomości w czasie. Główną ideą ML jest to, że komputer może zrobić to samo: może nauczyć się, jak wycenić dom na podstawie danych historycznych wcześniej sprzedanych, zasilając go zarówno ich cechami (rozmiar, sąsiedztwo itd.), jak i ostateczną sprzedażą Cena. Ta sztuczna inteligencja pozwoliłaby Ci mieć taką samą wiedzę fachową, jak doświadczony broker, przy tej samej szybkości, skali, dostępności i kosztach oprogramowania, umożliwiając każdemu sprzedającemu dom w ułamku sekundy oszacowanie wartości jego nieruchomości. To połączenie dostępności danych i wartości biznesowej jest świętym Graalem innowacji opartych na sztucznej inteligencji. Zidentyfikowaliśmy jasną propozycję wartości, popartą odpowiednim typem danych do jej zbudowania.

Dodanie możliwości AI do FutureHouse

Jednym z naszych celów jest nauczenie się, jak rozmawiać z naukowcami i inżynierami danych, więc jest to akapit, w którym zapoznasz się z żargonem i szczegółami technicznymi. Na tym etapie procesu tworzenia pomysłów zdecydowaliśmy, że chcemy zbudować automatyczny silnik prognozowania cen domowych. Wierzymy, że pomoże to zoptymalizować nasze wewnętrzne procesy i zaoferuje wartość naszym klientom, ponieważ będą mogli uzyskać wycenę swojego domu bezpośrednio online. Posiadamy również rejestr transakcji w regionie, w tym kilka ważnych cech domu (np. metraż) oraz cenę sprzedaży. Gdybyś przyniósł to wszystko inżynierowi, oto jak potoczyłaby się hipotetyczna rozmowa

Ty: Chcielibyśmy dodać automatyczny predyktor cen do naszej witryny z ofertami nieruchomości. Czy możesz nam w tym pomóc?

Inżynier: Fajnie - to standardowy problem w uczeniu maszynowym. Istnieje kilka dobrze znanych algorytmów uczenia maszynowego, co do których jestem przekonany, że mogą stworzyć dobry model. Jakie mamy dane wejściowe?

Ty: Rozmawialiśmy z agentami, którzy zazwyczaj biorą pod uwagę powierzchnię, liczbę łazienek, odległość od transportu publicznego i tym podobne. Oto pełna lista.

Inżynier: Jasne, lista funkcji wydaje się dobrym początkiem. A jakiego celu potrzebujesz od modelu? Tylko oczekiwana cena nieruchomości?

Ty: Zgadza się, to wszystko, czego potrzebujemy. Cena pomoże nam oferować bezpłatne wyceny w czasie rzeczywistym na naszej stronie internetowej.

Inżynier: Fajnie. Jakie masz na to dane?

Ty: Mamy zapisy dotyczące wszystkich transakcji domów w regionie w ciągu ostatniej dekady.

Inżynier: Świetnie! Czy te zapisy zawierają wszystkie funkcje, o których wspominałeś wcześniej?

Ty: Tak, większość z nich. Z pewnością dysponujemy powierzchnią i podobnymi rzeczami, ale być może będziesz musiał wymyślić coś dla transportu publicznego.

Inżynier: OK, mogę z tym pracować. Czy macie również etykiety dla nich wszystkich?

Ty: Tak, zapisy zawierają ostateczną cenę sprzedaży każdej nieruchomości.

Inżynier: Niesamowite. O ilu przykładach mówimy?

Ty: W ciągu ostatnich pięciu lat mamy około 200 000 transakcji.

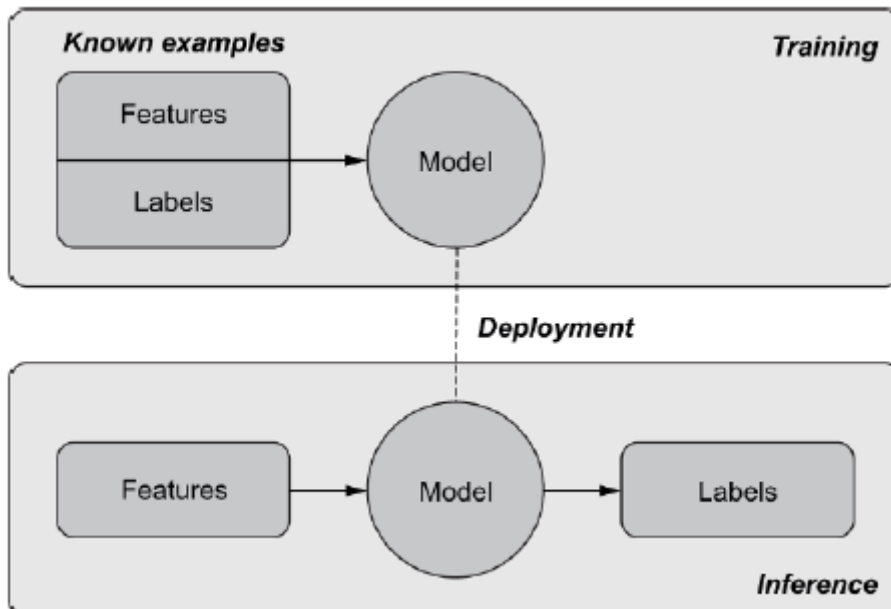
Inżynier: Fantastycznie! To powinno wystarczyć do wytrenowania naszego modelu.

Cofnijmy się trochę i przeformułujmy to, przez co właśnie przeszedłeś. Przede wszystkim inżynier określił Twój projekt jako problem uczenia maszynowego. Problem tekstowy w dziedzinie takiej jak ML nie ma negatywnej konotacji; jest używany w odniesieniu do dobrze zdefiniowanych zadań, które mają określony cel i znane warunki. Celem tego projektu jest zbudowanie modelu: zautomatyzowanego programu komputerowego, który może oszacować wartości nieruchomości dla naszego konkretnego obszaru. Nawet jeśli jego wewnętrzne działanie jest nieskończenie bardziej złożone, model jest jak kalkulator kieszonkowy: ma zestaw danych wejściowych, przetwarza je za pomocą algorytmu uczenia maszynowego, a następnie generuje cel, na którym nam zależy. Wiele osób lubi również mówić, że model przewiduje cel lub dokonuje prognozy (na przykład często czytasz o jakości prognoz modelu). W naszym przypadku dane wejściowe to zestaw funkcji opisujących każdy dom: powierzchnia kwadratowa, liczba pokoi i tak dalej. Inżynier pyta też o liczbę przykładów, czyli o liczbę sprzedanych domów w przeszłości, które mamy w naszych ewidencjach. Kiedy mówimy o oczekiwanej wartości wyjścia dla konkretnego przykładu (home), zwykle nazywamy to label, tak jakbyśmy oznaczali każdy przykład poprawną wartością wyjścia. Łatwo to zwizualizować w postaci arkusza Excela, takiego jak na rysunku

	A	B	C	D
1	Rooms	Square meters	Distance from city center	Price
2	2	80	5.4 km	100,000
3	1	42	7 km	80,000
4	3	120	23 km	160,000
5	2	65	2 km	70,000

Nasz inżynier ML poprosił o dane historyczne, ponieważ muszą poinstruować komputer, aby przeszedł przez proces zwany treningiem. Tak jak doświadczony pośrednik w handlu nieruchomościami może odgadnąć cenę nowego domu na podstawie swoich doświadczeń z domami sprzedanymi w przeszłości, algorytm uczenia maszynowego może nauczyć się przewidywać cenę domu na podstawie przykładów z przeszłości w fazie zwanej szkoleniem. W tej początkowej fazie inżynierowie piszą kod, który zasila model tysiącami przykładów, które posiadamy. Każdy przykład zawiera zarówno cechy (charakterystykę domu, na podstawie której należy się uczyć), jak i odpowiadającą im etykietę (cenę sprzedaży), dzięki czemu algorytm może uczyć się na podstawie wcześniejszych doświadczeń. Po zakończeniu szkolenia mamy samowystarczalny program komputerowy (model), który zawiera całą moc predykcyjną przewidywania ceny domu. Ponieważ model osadza teraz wiedzę zawartą w uczącym zestawie danych, możemy zacząć z niego korzystać niezależnie od uczących danych. Za każdym razem, gdy chcemy ocenić nowy dom, możemy wprowadzić jego cechy do modelu i uzyskać odpowiedź na temat jego prawdopodobnej ceny na rynku. Ten drugi krok, zwany wnioskowaniem, sprawia, że model jest przydatny dla organizacji, ponieważ generuje szacunki dotyczące nowych domów notowanych na naszej platformie. Gdy model zostanie przeszkolony, możemy poprosić inżynierów o podłączenie go do witryny FutureHouse i rozpoczęcie odpowiadania na prośby klientów, gdy przychodzą. To normalne, jeśli granica między szkoleniem a wnioskowaniem jest dla Ciebie nadal nieco rozmyta. W końcu ludzie chodzą przez swoje życie, jednocześnie szkoląc się na nowej wiedzy i wnioskując na temat przeszłych przekonań: nasz przyjaciel agent nieruchomości ocenia nieruchomości i uczy się od każdej sprzedanej w tym samym czasie od początku swojej kariery. Jednak ważne jest, abyś zdał sobie sprawę, że są to

dwa bardzo oddzielne kroki w świecie uczenia maszynowego. Gdy zestaw danych jest gotowy, szkolenie to niewiele więcej niż mechaniczne łączenie danych z różnymi modelami, które inżynier może zdecydować się przetestować. Prawdziwa wartość jest generowana podczas wnioskowania, kiedy możemy wykorzystać całe to doświadczenie z przeszłości, aby dokonać nowych przewidywań. To rozróżnienie między uczeniem a wnioskowaniem jest wyrażone graficznie na rysunku

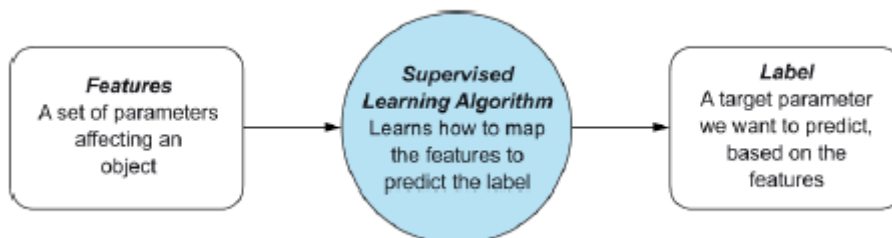


Do tej pory udało nam się przeformułować nasz problem prognozowania cen nieruchomości w sposób ogólny, co przyda się, gdy zajmiemy się innymi problemami. Prosimy inżynierów o zbudowanie modelu, który może mapować dane wejściowe (czyli cechy) do wartości docelowej (ceny domu). Przekażemy również pewne dane, w postaci długiej listy przykładów (przeszłe sprzedaże), każdy z pewnymi cechami (metraż kwadratowy i inne liczby) oraz etykietą (cena sprzedaży). Chcemy, aby model był w stanie przewidzieć cenę nowych domów, które pojawiają się na rynku, i robił to w sposób zgodny z podanymi przez nas przykładami. Innymi słowy, spodziewamy się, że model wyceni nowe dwupokojowe mieszkanie w nowojorskiej dzielnicy Tribeca, podobnie jak inne dwupokojowe sprzedane w pobliżu w zeszłym roku. Do tej pory wprowadziliśmy sporo abstrakcji, więc zatrzymajmy się na chwilę i przypomnijmy:

- * Algorytm uczenia maszynowego - technologia, która umożliwia komputerom uczenie się na podstawie danych
- * Cechy - zestaw cech obiektu, z których algorytm może się uczyć
- * Etykieta - wynik lub cel, który algorytm ma przewidywać
- * Uczenie - faza, w której algorytm uczenia maszynowego jest zasilany przykładami z przeszłości i uczy się z nich
- * Model - wynik fazy uczenia - samodzielny program komputerowy, który może przewidzieć etykietę, mając do dyspozycji zestaw funkcji
- * Wnioskowanie - faza, w której model jest używany z nowymi przykładami

Wspaniałą rzeczą w tych wszystkich abstrakcjach jest to, że możemy je teraz zastosować do innych problemów. Załóżmy na przykład, że chcemy przewidzieć cenę używanych samochodów, a nie domów.

Czy możesz pomyśleć, jakich funkcji chciałbyś użyć? Model samochodu i przebieg prawdopodobnie są dobrymi przypuszczeniami, podobnie jak historia konserwacji. Ale możesz mieć jeszcze więcej wyobraźni: pomyśl o dotacjach rządowych dla pojazdów o niskim poziomie emisji, liczbie wyszukiwań Google dla tych modeli i nie tylko. Mówiąc bardziej ogólnie, wszystkie problemy, dla których chcemy przewidzieć wartość (której jeszcze nie znamy) na podstawie zestawu informacji, które posiadamy, są przykładami nadzorowanego uczenia się. Uczenie nadzorowane to obszar uczenia maszynowego, który ma obecnie najwięcej zastosowań w przemyśle i badaniach naukowych. Nasz dotychczasowy przykład należy do tego podzbioru: wykorzystuje zestaw technik, które pozwalają maszynom uczyć się mapowaniem między zbiorem informacji zwanymi cechami a wartością docelową zwaną etykietą. Piękno nadzorowanego uczenia się i ogólnie uczenia maszynowego polega na tym, że komputer automatycznie znajdzie tę formułę, niezależnie od jej złożoności. Udany model zapewni prognozy zgodne z etykietami, przenosząc w ten sposób doświadczenie osadzone w zbiorze danych do nowych przypadków. Rysunek pokazuje, w jaki sposób funkcje i etykiety współdziałają w aplikacjach uczenia nadzorowanego.



Do tej pory poznałeś podstawową koncepcję uczenia nadzorowanego. Jak widać, to nie jest nauka o raketach. W rzeczywistości jest to łatwe do zrozumienia. Złożoność tkwi w wewnętrznych mechanizmach: algorytmach uczenia maszynowego, silnikach, które umożliwiają komputerom uczenie się. Nie będziemy zagłębiać się w te algorytmy, chyba że uznamy, że mogą być dla Ciebie interesujące i przydatne. Wierzmy, że dzięki dzisiejszym narzędziom i usługom możesz generować wartość dla swojej organizacji, wiedząc wystarczająco dużo, aby wymyślić produkty lub usługi AI i przekazać piłkę inżynierom ML, gdy nadejdzie na to czas. Teraz zaczynasz wiedzieć, jak zrobić jedno i drugie, więc przybij piątkę! Nawet jeśli poprzednie koncepcje wydają się proste, nie lekceważ swojej nowo odkrytej wiedzy: firmy, które wykorzystywały proste techniki, których się właśnie nauczyłeś, wyprodukowały miliardy dolarów. Dlaczego zatem uczenie maszynowe jest tak potężne? To jest temat następnej sekcji.

Zalety uczenia maszynowego

W poprzedniej części odbyliśmy przyjemną rozmowę z wymagowanym inżynierem i rozstaliśmy się wiedząc, że będzie budowanie modelu do przewidywania cen domów na podstawie niektórych jego funkcji, takich jak powierzchnia kwadratowa. Odrzuć na chwilę problem przewidywania ceny i wyobraź sobie, że potrzebujemy kalkulatora podatku od nieruchomości, który kupujący musiałby zapłacić. Gdyby tak było, nasza rozmowa z inżynierem byłaby znacznie krótsza:

Ty: Chcielibyśmy dodać do naszej strony kalkulator podatku od nieruchomości. Możesz pomóc?

Inżynier: Fajnie. Jaka jest odpowiednia sekcja w kodzie podatkowym?

Ty: Możesz odwołać się do 4123 i 564b.

Inżynier: Świetnie, to wszystko, czego potrzebuję.

Główna różnica między budową predyktora cen a kalkulatorem podatkowym polega na tym, że ten ostatni rządzi się precyzyjnymi regułami wyrażonymi już w formie matematycznej. Nawet najmłodszy księgowy może zastosować wzory matematyczne określające podatek od nieruchomości i zastosować je do nowych nieruchomości, bez potrzeby posiadania doświadczenia. To samo dotyczy fizyki: odkąd Isaac Newton odkrył prawa dynamiki w XVII wieku, byliśmy w stanie przewidywać ruchy ciał i budować rakiety, które wysyłały ludzi na Księżyc. Zawsze, gdy mamy takie sytuacje, matematyka i konwencjonalna informatyka wystarczą: wystarczy przetłumaczyć reguły i formuły na kod komputerowy, a wynik będzie w 100% dokładny. Inaczej jest w przypadku predykcji cenowej: związek między cechami domu a jego ceną sprzedaży jest niewyraźny, niejasny i zdecydowanie nie jest opisany w księdze zasad. Nawet najbardziej doświadczony agent nie jest w stanie spisać przyjaznych dla komputera reguł wyceny nieruchomości, a programista nie może przełożyć ich na oprogramowanie. W ubiegłym stuleciu ludzkość wykorzystywała komputery do realizacji niesamowitych celów, posuwając się nawet do wystrzeliwania rakiet w kosmos i przeprowadzania miliardów transakcji finansowych każdego dnia. A jednak nasze najbardziej wyrafinowane roboty nadal nie mogą przejść przez ulicę tak bezpiecznie, jak sześciolatnie dziecko. W rzeczywistości informatyka od dawna tkwi w wielu ważnych problemach, ponieważ były one po prostu zbyt złożone, aby można je było zrozumieć analitycznie. Tłumaczenie poezji na języki obce, prowadzenie samochodu czy odbieranie telefonów to zadania, które ludzie uczą się wykonywać poprzez doświadczenie, a nie poprzez wykonywanie instrukcji. Uczenie maszynowe w końcu umożliwiło komputerom robienie tego samego, całkowicie omijając reguły matematyczne, a tym samym umożliwiając zupełnie nowe obszary zastosowań. Podstawowym spostrzeżeniem jest to, że możemy wykorzystać doświadczenie, aby zrekompensować brak reguł i relacji matematycznych. Innymi słowami, model jest rozwijany poprzez poznanie odpowiednich informacji z przykładów z przeszłości. Jest to w przeciwieństwie do tradycyjnego programowania, w którym model jest opracowywany przez człowieka i mozolnie zakodowany w komputerze (czyli zaprogramowany). Podstawową ideą uczenia maszynowego jest to, że możemy wykorzystać funkcje i etykiety w danych, aby komputer samodzielnie nauczył się relacji między nimi. Po zapisaniu tych informacji celem jest, aby model replikował tę samą relację na niewidocznych przykładach (nowe domy), przewidując etykietę (cenę domu) przy danym zestawie cech (charakterystyka domu). Ogólna koncepcja nie różni się od tego, co zwykle robią ludzie: dzieci muszą być uczone różnicy między końmi a kucykami tylko kilka razy, zanim „zrozumieją” to i przyswoją różnicę. W porównaniu z konwencjonalną inżynierią oprogramowania uczenie maszynowe ma w zanadrzu dwa asy. Omówiliśmy już pierwsze: są pewne problemy, których po prostu nie możemy rozwiązać za pomocą komputera w żaden inny sposób. Drugim powodem korzystania z uczenia maszynowego jest to, że oprogramowanie jest statyczne: nawet jeśli inżynierowie potrafią wymyślić złożone reguły logiczne, nie mogą się one zmienić po dostarczeniu produktu. Dzięki uczeniu maszynowemu dodawanie nowych danych szkoleniowych może zarówno poprawić wydajność, jak i rozwiązać zupełnie nowe sytuacje, bez konieczności dodatkowego wysiłku inżynierskiego. Na przykład, gdy zostanie sprzedana fala nowych mieszkań, możemy po prostu powtórzyć szkolenie z nowymi danymi, a prognozy zostaną dostosowane.

Stosowanie AI do ogólnych podstawowych danych biznesowych

Zapewne można się domyślić, że sztuczną inteligencję można zastosować w większej liczbie branż niż w branży nieruchomości. Sztuczną inteligencję można zastosować w niezliczonych domenach. Ta sekcja pomoże Ci poszerzyć horyzonty poprzez uogólnienie podstawowych narzędzi wprowadzonych w tej części. Jak wspomnieliśmy wcześniej, nasz algorytm predykcyjny cen należy do gałęzi uczenia maszynowego zwanej uczeniem nadzorowanym. Nadzorowane algorytmy uczenia się oczywiście nie ograniczają się do wykrywania relacji między cechami domu a ich ceną sprzedaży. Pięknym atrybutem tych algorytmów jest to, że są to ogólne maszyny uczące się, które można przystosować do praktycznie każdego zjawiska, którego cechy wpływają na wartość etykiety. Przełączanie etykiet to najszybszy

sposób na docenienie elastyczności nadzorowanych algorytmów uczenia się. Załóżmy, że zastępujemy naszą obecną etykietę (cena sprzedaży) czasem sprzedaży domu. Ta metryka byłaby równie łatwa do uzyskania z naszych historycznych danych i byłaby przydatna dla klientów i agentów, aby mogli ocenić popyt na dom i odpowiednio działać. Wydaje się rozsądne, że duże, luksusowe domy sprzedawałyby się wolniej niż apartamenty z dwiema sypialniami. Nawet przy różnych etykietach moglibyśmy prawdopodobnie ponownie wykorzystać ten sam zestaw funkcji, które zidentyfikowaliśmy do przewidywania cen, aby otrzymać nowy algorytm, który zamiast przewidywać najlepszą cenę domu, będzie przewidywał czas potrzebny na znalezienie nabywcy. Jeśli za każdym razem, gdy otwierasz arkusz kalkulacyjny Excel, patrzysz na kolumny pod kątem potencjalnych funkcji i etykiet dla problemu z uczeniem maszynowym, otworzy się przed Tobą świat możliwości. Na przykład, jeśli pracujesz dla banku i jesteś przyzwyczajony do przeglądania arkuszy Excela z danymi klientów, możesz teraz zobaczyć, jak wykorzystać te dane jako funkcje i etykiety. Banki słyną z wykorzystywania nadzorowanej nauki do przewidywania ryzyka kredytowego swoich klientów przed udzieleniem kredytu, wykorzystując takie cechy, jak dochód, wiek i stan cywilny. Jeśli pracujesz dla firmy zajmującej się konserwacją i masz pewne dane o awariach sprzętu, nie powinieneś się dziwić, że niektóre firmy mogą przewidzieć prawdopodobieństwo awarii sprzętu w ciągu nowego miesiąca, korzystając z takich funkcji, jak jego wiek i ilość czasu, minęło od ostatniego remontu. Masz teraz całą wiedzę potrzebną do zrozumienia i sformułowania takich problemów, a także zdobędziesz więcej praktyki z analizami przypadków na końcu rozdziału. Wkrótce rozwinemy tę technologię na dwa główne sposoby: patrząc na wejścia lub wyjścia modelu. Jeśli chodzi o dane wejściowe, nauczymy Cię technik sztucznej inteligencji, które mogą wykorzystywać obrazy lub tekst jako dane wejściowe, znacznie rozszerzając rodzaje problemów, z którymi możemy sobie poradzić. Jeśli chodzi o wyniki, omówimy rodziny modeli, które nie tylko przewidują etykiety pasujące do danych szkoleniowych, ale wykonują inne przydatne rzeczy, takie jak grupowanie podobnych elementów lub pomaganie klientom w znajdowaniu interesujących treści. Nadszedł czas na podsumowanie technicznej części, więc przypomnijmy najważniejsze koncepcje. Dowiedziałeś się, że magia uczenia maszynowego polega na zastąpieniu analitycznego rozumienia złożonych procesów (takich jak rynek nieruchomości) modelem, który uczy się automatycznie na podstawie punktów danych. Dopóki przykłady dobrze odzwierciedlają zachowanie systemu, możemy wykorzystać wiedzę zawartą w modelu do przewidywania niewidzialnych zdarzeń (takich jak nowe domy na rynku). Widziałeś, że musimy wcześniej wykonać pewną pracę, aby zdefiniować wizję biznesową i zidentyfikować dane. Następnie jesteśmy gotowi poprosić inżynierów o przeszkolenie modelu. Po zakończeniu szkolenia otrzymujemy gotowy model, którego możemy użyć do wnioskowania, uzyskując w końcu pewną wartość biznesową. Podczas gdy ta i inne części skupiają się na tym, co możesz zrobić z AI, w częściach kolejnych powiemy więcej o tym, jak to urzeczywistnić. Teraz nadszedł czas, aby wrócić do tych dwóch studiów przypadku.

Jak Google wykorzystał sztuczną inteligencję do obniżenia rachunków za energię

W 2014 roku Jim Gao był inżynierem w Google. Odpowiadał za to, aby ogromne systemy klimatyzacji w centrach danych (DC) giganta technologicznego działały tak płynnie i wydajnie, jak to tylko możliwe. Mając doświadczenie w inżynierii mechanicznej, Gao postępował zgodnie z najlepszymi praktykami w branży energetycznej i osiągał świetne wyniki. Po wdrożeniu najczęściej stosowanych środków oszczędzania energii, wydajność DC Google zaczęła spadać, odstawiając ograniczenia tradycyjnego podejścia do oszczędzania energii. Było jasne, że potrzebne jest nowe podejście. Gao postanowił podążać niepokonaną ścieżką, korzystając z polityki Google 20% - inicjatywy, która pozwala pracownikom spędzać 20% czasu na pracy nad tym, co ich zdaniem przyniesie największe korzyści Google. Jako inżynier centrum danych doskonale zdawał sobie sprawę z czujników rozmieszczonych w centrach danych i dużej ilości danych zbieranych z nich do celów operacyjnych. Gao postanowił zająć

się uczeniem maszynowym i spróbował zbudować modele do przewidywania i poprawy wydajności DC.

PROBLEM ZUŻYCIA ENERGII W CENTRUM DANYCH

Centrum danych to budynek, w którym znajdują się serwery sieciowe. W przypadku Google maszyny te obsługują zapytania w wyszukiwarce i mapach, przechowują zdjęcia i dokumenty oraz wykonują wszystkie inne zadania, których potrzebuje Google, aby oferować swoje usługi użytkownikom. Zużycie energii jest głównym czynnikiem generującym koszty DC ze względu na dużą liczbę energochłonnych komputerów, w których się znajdują. Wysokie koszty nie są jedynym czynnikiem, który należy wziąć pod uwagę podczas ważenia zużycia energii DC; ważny jest również wpływ na środowisko. DC zużywają obecnie 2% światowej energii elektrycznej, a liczba ta z pewnością wzrośnie wraz ze wzrostem zapotrzebowania na usługi sieciowe. Ilość energii zużywanej do zasilania komputerów nie może być łatwo zoptymalizowana przez zespół operacyjny, ponieważ zależy to od obciążenia obliczeniowego i wydajności chipów. Z tego powodu inżynierowie centrów danych dążą do zmniejszenia dodatkowego zużycia. Wydajność centrów danych jest zwykle mierzona poprzez śledzenie miernika zwanego efektywnością zużycia energii (PUE). Ta metryka odzwierciedla, ile energii jest zużywane na cokolwiek innym niż rzeczywiste komputery, które tworzą centrum danych:

$PUE = \text{Całkowita energia obiektu} / \text{Energia sprzętu IT}$

Idealne centrum danych ma PUE 1: cała energia jest zużywana na zasilanie komputerów. Im wyższy PUE, tym więcej energii zużywa się na inne systemy, wśród których najważniejsze jest chłodzenie. Na przykład PUE 1,5 oznacza, że na każdą kilowatogodzinę (kWh) energii zużywanej do zasilania komputerów potrzebne jest dodatkowe 0,5 kWh energii na chłodzenie i inne drobne potrzeby. Google zawsze był liderem efektywności PUE. Według badania przeprowadzonego przez Uptime Institute w 2018 r. wśród 900 operatorów DC, średnia PUE w branży wyniosła 1,58. Google stale ulepszał swój PUE, aż osiągnął 1,12 w 2013 roku. Niestety ta wartość poprawiła się dopiero w 2017 roku.

PODEJŚCIE ML DO WYDAJNOŚCI CENTRUM DANYCH

Gao zdał sobie sprawę, że jedną z przeszkód w dalszym obniżaniu PUE jest to, że bardzo skomplikowane jest prawidłowe przewidzenie go w różnych scenariuszach przy użyciu tradycyjnego podejścia inżynierskiego, ze względu na złożone interakcje między czynnikami (na przykład wiatr może pomóc schłodzić elektrownię i zmniejszyć potrzeba sztucznego chłodzenia). Z drugiej strony doskonale zdawał sobie sprawę z dużych zbiorów danych gromadzonych przez jego zespół w ramach codziennych operacji dzięki tysiącom czujników rozmieszczonych w komponentach, które zbierają miliony punktów danych. Gao przewidział potencjał wykorzystania tych danych do wytrenowania modelu ML zdolnego do przewyższenia ograniczeń tradycyjnej termodynamiki. Jego pierwszym podejściem było zbudowanie prostej sieci neuronowej (klasycznego algorytmu używanego do budowania modeli uczenia nadzorowanego), która została wytrenowana w celu przewidywania PUE, mając na uwadze listę cech, które na niego wpływają. Gao używało łącznie 19 funkcji, w tym:

- * Całkowite obciążenie IT serwera
- * Całkowita liczba uruchomionych pomp wody procesowej
- * Średnia nastawa temperatury wody wypływającej z chłodni kominowej
- * Całkowita liczba pracujących agregatów chłodniczych
- * Średnia temperatura zbliżenia wymiennika ciepła

* Temperatura termometru mokrego i termometru suchego na zewnątrz

* Wilgotność względna powietrza zewnętrznego, prędkość i kierunek wiatru

Gao wykorzystał 184 435 próbek czasu w rozdzielczości pięciominutowej (około dwóch lat danych operacyjnych). Ostateczny model był w stanie przewidzieć DC PUE w zakresie 0,004 +/- 0,0005, błąd około 0,4%, dla PUE 1,1. Gao był w stanie szybko zbudować pierwszy model weryfikacji koncepcji (POC) przy użyciu frameworków kodowania open source. Ostateczny wynikowy model został wykorzystany do trzech głównych zastosowań:

* Automatyczne ostrzeżenie o wydajności i rozwiązywanie problemów poprzez porównanie rzeczywistej i przewidywanej wydajności DC dla dowolnego zestawu warunków

* Ocena wrażliwości PUE na parametry operacyjne

* Uruchamianie symulacji cyfrowych w różnych konfiguracjach bez wprowadzania zmian fizycznych

Wyniki pracy Gao miały duży wpływ na firmę i zostały publicznie docenione przez wiceprezesa centrum danych, Joe Kava, który opisał pracę i wyniki w oficjalnym poście na blogu Google. „Better Data Centers Through Machine Learning” podkreślił, w jaki sposób modele Gao były w stanie zidentyfikować wzorce danych, których dana osoba nie może dostrzec, prowadząc do modelu zdolnego do przewidywania PUE z 99,6% dokładnością. Gao i jego zespół zaczęli używać tego modelu, aby wymyślić nowe sposoby poprawy wydajności. Kiedy serwery były odłączane, wiadomo było, że powodowało to na przykład niższą wydajność w centrum danych. Dzięki modelom Gao zespół Google DC był w stanie przeprowadzić symulacje zachowania centrum danych i znaleźć nowe sposoby ograniczania utraty wydajności, oszczędzania energii i pieniędzy.

ML W ZASTOSOWANIU DO OPTIMALIZACJI CENTRUM DANYCH NABIERA ROZPĘDU

Praca Gao przyciągnęła uwagę na wyższych szczeblach firmy. Został awansowany na kierownictwo wielofunkcyjnego zespołu inżynierów zajmujących się uczeniem maszynowym, oprogramowaniem, elektrykami, mechanikami, sterowaniem i operacjami w celu opracowania kompleksowego rozwiązania inteligentnego centrum danych. Doświadczenie ML w nowym zespole zostało wniesione przez DeepMind, brytyjską firmę specjalizującą się w najnowocześniejszych algorytmach sztucznej inteligencji, która została przejęta przez Google w 2014 roku za 500 milionów dolarów. Dodatkowa wiedza fachowa w zakresie ML okazała się kluczowym atutem w osiągnięciu wydajności niespotykanej wcześniej w historii firmy, zmniejszając koszty ogólne PUE o 15% i powodując 40% ogólne obniżenie rachunków za energię. Nawet jeśli model DeepMind był bardziej złożony i dokładny niż podejście Gao i pozwolił Google obniżyć rachunki za energię o 40%, nadal generował zalecenia, które musiały zostać sprawdzone i wdrożone przez człowieka. Kolejnym logicznym krokiem projektu było całkowite przeniesienie zarządzania zasilaniem DC na rozwiązanie sztucznej inteligencji, zachowując jako nadzorców ludzkich ekspertów. Modele DeepMind zostały wdrożone w sierpniu 2017 r. i już przyniosły 10% oszczędności energii. Podczas gdy nowe modele działały, generowały dodatkowe dane, które zostały wykorzystane do ponownego szkolenia modeli i dalszej poprawy ich wydajności. W ciągu roku oszczędności energii z 10% wzrosły do 30%. System sterowania AI wykorzystywał nowatorskie sposoby zarządzania chłodzeniem, które zapewniały najwyższą wydajność w sposób nieoczekiwany i niezbadany przez doświadczonych operatorów DC. Na przykład Dan Fuenffinger, jeden z operatorów centrów danych Google, zauważył, że był zaskoczony, gdy system AI był w stanie samodzielnie wykorzystać warunki zimowe i produkować wodę o niższej temperaturze, zmniejszając zużycie energii.

PYTANIA DOTYCZĄCE PRZYPADKU

1. Poprawa PUE centrum danych jest trudnym problemem i Gao wyczuł, że ML może być rozwiązaniem tego problemu. Jakie są cechy tego wyzwania, które sprawiają, że dobrze pasuje do ML?
2. Jakie są kluczowe elementy, które pozwoliły Gao samodzielnie zbudować pierwszy POC?
3. Jak Google zmienił inicjatywę jednego inżyniera w sukces całej firmy?

OMÓWIENIE PRZYPADKU

Projekt centrum danych Google to klasyczny przypadek optymalizacji, do którego ML jest idealnym rozwiązaniem. Rozpoczyna się od kluczowych wskaźników wydajności (KPI), których przewidywanie leży w interesie biznesowym (PUE) oraz dużego zestawu parametrów, które na nie wpływają (temperatury, ciśnienia, obciążenie itd.). Inżynierowie próbowali zbudować model, który może odwzorować te parametry na PUE, ale tradycyjne metody, które opierają się na termodynamice, mają poważne wady, które uniemożliwiły im dostarczanie wyników powyżej pewnego progu. Tak powinien wyglądać problem biznesowy uczenia maszynowego: złożona, nieznana relacja między zbiorem zmiennych, które wpływają na metrykę, która ma bezpośredni wpływ na biznes. Gao wykonał świetną robotę, dostrzegając te cechy i zdając sobie sprawę, że ML może być właściwym młotkiem do uderzenia w tę igłę. Kilka elementów pozwoliło Gao odnieść sukces. Pierwszym z nich jest dostępność danych, bez których realizacja rozwiązania ML nie byłaby możliwa. Gao wiedział ze swojego doświadczenia zawodowego o istnieniu tego zestawu danych i był w stanie, dzięki sprawnemu zarządzaniu danymi w swojej firmie, łatwo pobrać dane, których potrzebował do rozpoczęcia eksperymentów. Później omówimy bardziej szczegółowo zarządzanie danymi, ale jasne jest, że dostępność i dostęp do danych ma kluczowe znaczenie dla rozpoczęcia pracy nad takim projektem ML. Zauważ, że te dane były dostępne od lat, ale nikt nie wyszedł poza ich zwykły cel operacyjny. Innym czynnikiem, który pomógł Gao w rozpoczęciu pracy, była dostępność narzędzi open source, które umożliwiły mu szybkie zbudowanie POC. To pozwoliło mu zdobyć uwagę i wsparcie korporacyjne, których potrzebował, aby przenieść swoje podejście na wyższy poziom, wprowadzając na pokład wykwalifikowanych inżynierów ML z DeepMind i opierając się na swojej pierwszej intuicji. Wreszcie sposób, w jaki ten projekt został zrealizowany, jest doskonałym przykładem świetnej strategii ML. Omówimy ten temat bardziej szczegółowo w drugiej części, ale zauważ, że projekt nie rozpoczął się od wielkiej misji i złożonego rozwiązania. Pierwszy krok był na tyle prosty, że mógł go zbudować jeden inżynier w 20% swojego czasu, ale także wystarczająco mocny, aby udowodnić potencjał tego podejścia. Drugi krok był bardziej złożony, a osiągnięcie wyższej wydajności wymagało zdobycia nowej wiedzy, wykonania testów, nabrania rozpędu i przejścia do kolejnego etapu całkowicie zautomatyzowanej kontroli AI. Google zaplanowało, że projekt będzie budowany etapami, z których każdy ma wymierny zwrot z inwestycji (ROI) i opiera się na poprzednim, aż do osiągnięcia najwyższej wydajności i dużego wpływu na wynik finansowy. Podsumowując, to studium przypadku ma następujące główne wnioski:

* Problem, który może dobrze pasować do ML, musi mieć pewne cechy:

- Istnieje ważny KPI.
- Na KPI wpływa szereg mierzalnych parametrów.
- Parametry wpływają na KPI poprzez złożoną i nieznaną zależność.

* Operacyjne zbiory danych mogą mieć ogromną ukrytą wartość.

* Technologie open source mogą umożliwić małym zespołom inżynierów tworzenie POC, które przynoszą wymierne rezultaty dla firmy i udowadniają, że koncepcja może być skalowana i dalej ulepszana.

* Podczas tworzenia projektów ML zawsze dobrze jest zacząć od małych, identyfikując najmniejsze zadanie, które może dać natychmiastowy efekt biznesowy, aby zbudować podstawę dla bardziej złożonych projektów.

Jak firma Square wykorzystała sztuczną inteligencję do pożyczania miliardów małym firmom?

To studium przypadku dotyczy tego, jak startup fintech rzucił wyzwanie dużym bankom, pożyczając kapitał tradycyjnie niedocenianym małym firmom. Wykorzystując uczenie maszynowe w swoim zbiorze danych dotyczących transakcji detalicznych, Square Capital może proaktywnie oferować małe pożyczki firmom, zanim nawet o nie poproszą. Wskaźnik przestępczości Square wynosi połowę średniej w branży.

SQUARE I SQUARE CAPITAL

Square to firma świadcząca usługi finansowe skoncentrowana na potrzebach małych firm. Została założona przez założyciela Twittera Jacka Dorsey'a w 2009 roku i weszła na giełdę w 2015 roku. Firma z siedzibą w San Francisco miała w 2019 roku kapitalizację rynkową 26 miliardów dolarów i zatrudnia 2300 osób. Pierwszym produktem Square był czytnik kart kredytowych podłączany bezpośrednio do smartfona, aby umożliwić małym sprzedawcom i profesjonalistom przyjmowanie płatności kartą kredytową. Sprzęt czytnika był dystrybuowany za darmo, bez konieczności posiadania miesięcznego abonamentu, co dało początek reputacji Square wśród małych sprzedawców detalicznych. Kultowy projekt sprzętu i stały przepływ rekomendacji również przyczyniły się do cementowania reputacji Square (do tej pory ekspansja międzynarodowa była ograniczona). W pierwszych latach działalności Square koncentrował się na niszy małych firm o obrotach poniżej 125 000 USD; przedsiębiorstwa te były źle obsługiwane przez banki i konwencjonalne podmioty przetwarzające płatności. Niskie opłaty i łatwa w obsłudze strona internetowa Square pomogły wciągnąć przedsiębiorców do tego, co wyrastało na kompletny ekosystem finansowy. W latach poprzedzających wprowadzenie na giełdę Square uzupełnił swoje usługi, uruchamiając aplikację do płatności peer-to-peer, platformę do zarządzania relacjami z klientami (CRM) i marketing oraz sprzęt do punktów sprzedaży dla większych sprzedawców. W 2014 roku Square uruchomił Square Capital, kolejny element swojego ekosystemu finansowego. W miarę rozwoju małych firm często pojawiają się problemy z płynnością finansową, które można złagodzić dzięki dodatkowemu kapitałowi obrotowemu. Jednak konwencjonalne banki często nie są wrażliwe na potrzeby tych klientów, ponieważ małe kwoty nie uzasadniają narzutów związanych z marketingiem i obsługą pożyczki. Square Capital wykorzystuje dane dotyczące transakcji w punktach sprzedaży, aby proaktywnie oferować swoim klientom niewielkie pożyczki. Dla klientów doświadczenie jest bezproblemowe, ponieważ mogą przyjmować pożyczkę i zarządzać nią bezpośrednio z tej samej witryny zarządzania, z której korzystają w przypadku innych usług w ekosystemie. Pożyczki nie mają określonego harmonogramu spłat: środki są automatycznie potrącone z transakcji kartą kredytową przetwarzanych przez Square. Połączenie bezproblemowego doświadczenia i istniejącej bazy klientów pozwoliło Square Capital na szybki rozwój: 200 000 sprzedawców pożyczyciło ponad 3,1 miliarda dolarów za pośrednictwem platformy tylko w latach 2014-2018. Podczas gdy modele ryzyka kredytowego są podstawą branży finansowej od wielu dziesięcioleci, Square ma dostęp do znacznie obszerniejszego zbioru danych, w tym sezonowości i czasu każdego zakupu. Dzięki temu Square może budować modele z doskonałą widocznością pozycji przepływów pieniężnych każdej firmy i odpowiednio dobierać wielkość oferowanej pożyczki. Wraz z prostymi opcjami spłaty, obniża to wskaźnik przeterminowanych pożyczek do 4%, czyli do połowy średniej w branży.

KRYTYKA I KONKURENCJA

Z finansowego punktu widzenia pożyczki oferowane przez Square Capital nie są niczym nowym. Pożyczkodawcy od niepamiętnych czasów oferowali kupcom zaliczki gotówkowe (MCA) na pokrycie

krótkoterminowych potrzeb gotówkowych firm. W porównaniu z MCA, które nie mają określonego terminu, pożyczki Square Capital muszą zostać spłacone w ciągu 18 miesięcy. Dzięki temu Square może przekazywać pożyczki szerszemu gronu instytucji finansowych, które mają trudności z obsługą MCA bez znanej daty spłaty. Square działa jedynie jako inicjator pożyczek, a nie pożyczkodawca, utrzymując atrakcyjny bilans porównywalny z innymi firmami technologicznymi. Podczas gdy większość instytucji finansowych wykorzystuje uczenie maszynowe do oceny ryzyka kredytowego, obecne zbiory firm fintech, takich jak Square, również zostały skrytykowane za usunięcie zbyt dużego nadzoru ludzkiego z procesów składania wniosków kredytowych. Na przykład firma Prosper zajmująca się pożyczkami peer-to-peer została poddana analizie w 2015 r. po tym, jak pożyczyła 25 000 USD parze terrorystów, która stała za strzelaniną w San Bernardino (Kalifornia), w której zginęły 24 osoby. Podczas gdy społeczność technologiczna jest podekscytowana tym, że fintech jest w stanie zakłócić i skalować tradycyjną bankowość za pomocą uczenia maszynowego, niektóre społeczności nie są tak optymistyczne. Na przykład te same małe firmy, które cieszą się szybkim czasem realizacji pożyczek Square Capital, również głośno mówią o postrzeganej nieuczciwości i braku przejrzystości, gdy ich wnioski o pożyczkę zostają odrzucone. Ponieważ Square nie publikuje żadnych szczegółów na temat swoich modeli oceny ryzyka, niektórzy handlowcy są zmuszeni do odtworzenia algorytmicznych decyzji. Z technicznego punktu widzenia jest to związane zarówno z dokładnością (wyrażoną w kategoriach prawda/falszywie dodatnie), jak i stronniczością modelu. Dzięki integracji z ekosystemem Square Square Capital może oferować atrakcyjne funkcje w porównaniu z tradycyjnymi bankami i MCA, które mają mniej bezpośredni dostęp do przepływu transakcji sprzedawcy. Jednak inne usługi, takie jak PayPal Working Capital, również mają dostęp do tego źródła danych, co utrudnia wyróżnienie Square Capital. Ankiety przeprowadzone wśród właścicieli małych firm sugerują, że zazwyczaj wybierają oni pożyczkodawcę na podstawie postrzeganej szansy na zatwierdzenie, a nie kwot lub warunków pożyczki. Oznacza to, że Square Capital może być zmuszone skupić się na doświadczeniu użytkownika i integracji z resztą ekosystemu, zamiast na optymalizacji modeli oceny ryzyka.

PYTANIA DOTYCZĄCE PRZYPADKU

1. Ryzyko kredytowe to dobrze poznana aplikacja uczenia maszynowego. Jak myślisz, co Square wypadło dobrze w porównaniu z przeszłą historią i obecną konkurencją?
2. Jakie lekcje możesz wyciągnąć z interakcji między uczeniem maszynowym a organizacjami regulacyjnymi (w tym przypadku bankowymi)?
3. Czy możesz opracować strategię zbierania danych treningowych dla tej aplikacji? Jak zmierzyłbyś wydajność powstałych modeli?

DYSKUSJA PRZYPADKU

Przedstawiliśmy to studium przypadku dotyczące Square Capital, ponieważ jest to historia sukcesu integracji modelu biznesowego opartego na sztucznej inteligencji z istniejącym ekosystemem usług. Ponadto omawia również problemy związane z wprowadzaniem funkcji opartych na ML do silnie regulowanej branży, takiej jak bankowość. Podstawowym produktem Square był historycznie system punktów sprzedaży dla małych firm. Jego mocnymi stronami były projekt, wrażenia użytkownika i obsługa klienta, a nie najnowocześniejsza technologia. Jednak dane były w rzeczywistości podstawą działalności Square, ponieważ wykrywanie oszustw i zapobieganie im są podstawowymi aspektami każdej firmy świadczącej usługi finansowe. To pokazuje, że gdy organizacja ustanowi potężną kulturę wokół danych (w tym zbieranie, przechowywanie i udostępnianie ich pracownikom), naturalne jest wprowadzanie na jej podstawie nowych produktów. Gdy istnieje dobrze zaprojektowana infrastruktura danych, modele ryzyka kredytowego można zintegrować z tą samą platformą, która przeszukiwała kanały transakcyjne w czasie rzeczywistym pod kątem oszustw, dodając wartość do tych

inwestycji w podstawową technologię. Z punktu widzenia doświadczenia użytkownika Square Capital jest dobrym przykładem bezproblemowego wprowadzania funkcji opartych na sztucznej inteligencji do istniejącej platformy w sposób całkowicie przejrzysty i naturalny dla użytkowników końcowych. Opierając się na zaufaniu, jakie kupcy już pokładają w Square, oferty Square Capital są skromne i odpowiednie dla odbiorców. Nie ma niepotrzebnego przechwalania się „oceną ryzyka w oparciu o sztuczną inteligencję” lub niejasnym marketingiem opartym na technologii. Pozostało tylko skupić się na doświadczeniu klienta. Jeśli chodzi o finanse przedsiębiorstw, warto zobaczyć, jak elastyczność, jaką Square uzyskał dzięki wykorzystaniu ML do oceny ryzyka kredytowego, pozwoliła mu na eksperymentowanie z nowymi modelami biznesowymi. Oferowanie pożyczek na czas określony zamiast tradycyjnych MCA umożliwiło Square skoncentrowanie się na swojej podstawowej misji, a nie na zaciąganiu samych pożyczek. Dzięki temu Square może osiągnąć gwałtowny (około 50% rocznie) wzrost, który wykorzystuje swoje mocne strony bez obciążania bilansu dodatkowym ryzykiem. Ciekawym elementem tego przypadku są wady ML w konkretnych przypadkach narożnych. Autonomiczne uczenie się algorytmów uczenia maszynowego uznaliśmy za mocną stronę, ale czasami przy podejmowaniu decyzji mamy trudności z zajrzeniem do logiki, za którą podążają algorytmy. Jeśli wniosek o pożyczkę małej firmy zostanie odrzucony przez algorytm ML, wyjaśnienie, dlaczego może być trudne lub niemożliwe. Ten aspekt uczenia maszynowego omówimy bardziej szczegółowo w ostatnim rozdziale. Square zaczął od silnej pozycji, ponieważ jej model biznesowy pozwalał firmie gromadzić podstawowe dane biznesowe dotyczące wszystkich klientów: ich historię transakcji. Ten zasób danych jest filarem, na którym zbudowany jest Square Capital. Zwróć uwagę, że te dane zostały zebrane w sposób naturalny, aby zaoferować standardową usługę POS firmy Square, i stały się niesamowitym atutem później, gdy wprowadzono Square Capital.

Lekcje ze studiów przypadku

Zarówno Google, jak i Square stworzyły aplikacje AI, które dawały im dużą przewagę nad konkurencją, a jednocześnie wyróżniały się u podstaw. Google wykorzystał sztuczną inteligencję do obniżenia kosztów i usprawnienia operacji, podczas gdy Square wykorzystał ją do uruchomienia zupełnie nowej usługi. Podstawowa zasada pozostaje taka sama: zarówno Square, jak i Google wykorzystywały sztuczną inteligencję do podstawowych danych biznesowych, tworząc ogromną wartość z tego, co większość innych firm postrzega jako koszty utopione. Zwróć uwagę, że zarówno Google, jak i Square korzystały z danych, które pierwotnie nie były zbierane z myślą o sztucznej inteligencji. W przypadku Google powszechną praktyką jest przechowywanie danych operacyjnych z zakładów klimatyzacyjnych centrum danych, głównie w celu raportowania i monitorowania, czy systemy działają zgodnie z oczekiwaniami. Pomysł Jima Gao, by wykorzystać ML do poprawy wydajności systemu, był nowatorski i pokazuje świetną intuicję inżyniera Google. To samo dotyczy Square: zbierał dane o transakcjach klientów w celu oferowania usług, takich jak analityka sprzedaży, ale oferowanie pożyczek nie było pierwotnym powodem gromadzenia tych danych. Przede wszystkim Square nie oferował nawet usługi pożyczki dla firm: jest to całkowicie nowa usługa udostępniana przez ML i ten bogaty zbiór danych, który posiadał. Umożliwiło to Square wejście na nowy rynek, który był całkowicie niedoceniany, ponieważ tradycyjne banki nie mają zasobów do oceny kwalifikowalności kredytowej małych firm. Innym aspektem, o którym należy pamiętać z tych dwóch studiów przypadku, jest to, że w obu stosowano metody nadzorowanego uczenia się. Google nauczył się przewidywać PUE zakładu chłodniczego na podstawie jego parametrów operacyjnych, a Square określił kwalifikowalność małej firmy na podstawie jej historii transakcji. W przypadku Google uczenie nadzorowane było szczególnie przydatne w przezwyciężaniu ograniczeń tradycyjnych rozwiązań problemu efektywności energetycznej. Aplikacja Square została wykorzystana do zautomatyzowania procesu, który byłby zbyt czasochłonny, gdyby został stworzony przez ludzi.

Ocena wydajności i ryzyka

Teraz czas na spotkanie z kierownictwem FutureHouse, aby usiąść i porozmawiać o ryzykach i obawach realizowanego przez nas projektu. Znaleźliśmy świetne uzasadnienie biznesowe dla algorytmu przewidywania wartości nieruchomości, zebraliśmy właściwe dane i zbudowaliśmy je. Teraz wystarczy tylko nacisnąć czerwony przycisk, dzięki któremu funkcja pojawi się na stronie dla użytkowników. Masz zamiar iść dalej; wszyscy w biurze patrzą na ciebie. Twoje ręce są spocone i roztrzęsione. Wykonujesz skok wiary, naciskasz przycisk, zamykasz oczy i masz nadzieję, że algorytm zadziała. Na szczęście nie tak dzieje się w prawdziwym świecie. Wdrażanie modeli bez dokładnego pomiaru ich wydajności jest nieodpowiedzialne. Po pierwsze: uczenie maszynowe opiera się na danych i statystykach, a zatem jego przewidywania prawdopodobnie nigdy nie będą w 100% dokładne. Chcemy jednak, aby prognozy były wartościowe dla biznesu, przy jednoczesnym poszanowaniu wszelkich standardów bezpieczeństwa i etyki. Ogólne systemy, które budujemy, muszą być odporne na ten margines błędu z tego samego powodu, dla którego używamy uczenia maszynowego: niektóre z tych problemów są tak złożone, że żadna ilość danych nie jest w stanie ich wszystkich uchwycić. Firmy używające ML w sytuacjach życia lub śmierci (takich jak samochody autonomiczne) muszą być o wiele bardziej ostrożne i konserwatywne niż marketerzy używający go do optymalizacji przekazu do klientów. Na szczęście wszystkie aplikacje, o których mówimy w tej książce, należą do tej drugiej kategorii. Ocena wydajności modelu ma dwa główne aspekty. Dobry sposób myślenia o nich może być znany z literatury dotyczącej zarządzania: znane niewiadome i nieznanne niewiadome. Główną znaną niewiadomą jest dokładność: miara tego, jak dobry model jest w dopasowywaniu cech do etykiet. W problemie przewidywania wartości domu przekłada się to na to, jak blisko przewidywanej ceny jest cena rzeczywista po sprzedaży domu. W przypadku problemów z uczeniem nadzorowanym istnieje sprytny sposób pomiaru dokładności modelu. Inżynierowie zwykle losowo dzielą dane treningowe na dwie części (powiedzmy 70% i 30% danych), jak widać na rysunku 2.8. Trenują model tylko na pierwszej części, a drugą część wykorzystują jako zestaw testowy do symulacji rzeczywistego scenariusza modelu działającego na nieznanymi danych. Ponieważ zestaw testowy był również częścią początkowego zestawu danych, mamy dla niego etykiety. Dlatego możemy porównać przewidywane etykiety z rzeczywistymi i zmierzyć dokładność modelu w sytuacji, która dość dobrze pasuje do świata rzeczywistego. Odłożenie zestawu testowego na bok jest ważne, ponieważ testowanie danych, które zostały użyte do szkolenia, przypomina trochę egzaminowanie uczniów z pytaniami, na które już widzieli rozwiązania. W przypadku przewidywania ceny domu można sobie wyobrazić wybranie jednego lub dwóch domów w każdym bloku, usunięcie ich ze zbioru danych używanego do uczenia i dodanie ich do zestawu testowego. To da nam dobre wyobrażenie o tym, jak dobrze działa model. Jeśli chodzi o wykorzystanie sztucznej inteligencji w organizacjach, dokładność nie jest miarą sukcesu. Niedokładny model zastosowany do właściwego problemu może być o wiele bardziej wartościowy dla firmy niż bardzo dokładny model, który rozwiązuje nieistotny problem. Złotą zasadą, o której należy pamiętać, jest statystyk George Box:

Wszystkie modele są błędne, ale niektóre są przydatne.

Jeśli dokładność jest dobrym miernikiem do pomiaru znanych niewiadomych, co z (bardziej przerażającymi) nieznanymi niewiadomymi? W takich sytuacjach model daje całkowicie niepoprawną lub bezsensowną odpowiedź. Najczęściej dzieje się tak dlatego, że dane treningowe były niekompletne i próbujemy przewidzieć cel w sytuacji (kombinacji cech), której nigdy wcześniej nie widziano podczas treningu. Nawet jeśli wnioskowanie w niewidocznych sytuacjach jest celem korzystania z uczenia maszynowego, całkowicie nieprzewidywalne zapytania mogą nadal zepsuć model. Na przykład nieprawidłowe wprowadzenie danych może spowodować, że model będzie próbował odgadnąć cenę domu z 50 łazienkami w przypadku, gdy w zestawie danych szkoleniowych w ogóle nie było takiej rezydencji. Większość rodzin modeli nie sygnalizuje ostrzeżenia w takich sytuacjach, więc inżynierowie

muszą wprowadzić pewne sprawdzanie poprawności w wejściach i wyjściach, aby zapewnić bezpieczeństwo. Inny ważny czynnik ryzyka dotyczy sedna tego, co sprawia, że uczenie maszynowe jest tak wspaniałe: automatycznego wyodrębniania mapowania między celem a funkcjami. Jest to potężna umiejętność, ale oznacza również, że inżynierowie nie mogą łatwo zrozumieć, dlaczego model dokonuje określonej prognozy, biorąc pod uwagę zestaw funkcji. Właśnie dlatego uczenie maszynowe jest często nazywane czarną skrzynką: wiesz, co wchodzi (funkcje), a co wychodzi (cel), ale trudno jest zrozumieć, dlaczego model wybrał te dane wyjściowe. Podczas gdy niektóre rodziny modeli są lepsze od innych, prawda jest taka, że jest to jeden z najbardziej niepokojących aspektów związanych z wykorzystaniem sztucznej inteligencji na dużą skalę. Nie obawiajcie się, bo temu zagadnieniu poświęcamy sporo materiału w drugiej części książki. Wiele omówiliśmy w tym rozdziale. Wprowadziliśmy Cię w główne koncepcje nadzorowanego uczenia się, pracując nad projektowaniem funkcji sztucznej inteligencji skierowanej do klienta: przewidywanie wartości nieruchomości dla fikcyjnej firmy pośrednictwa w obrocie nieruchomościami. Dowiedziałeś się, jaką wartość mają dane i jak są do tego przyzwyczajone aby zbudować model uczenia maszynowego. Studia przypadków pomogły Ci zastosować tę wiedzę w sytuacjach branżowych, dzięki czemu zaczniesz rozpoznawać podobne możliwości we własnej organizacji. Na zakończenie ciekawostka. Zillow to wiodący internetowy rynek nieruchomości w Stanach Zjednoczonych. Przewidywanie wartości właściwości było dla Zillow tak ważne, a uczenie maszynowe tak oczywistym rozwiązaniem, że zorganizowano konkurs na najlepszy algorytm. Zwycięska nagroda? Milion dolarów.

Podsumowanie

- * Podstawowe dane biznesowe to rodzaj danych, które są najsilniej powiązane z silnikiem generującym wartość organizacji. Dlatego ma również największy potencjał oddziaływania AI.
- * Uczenie nadzorowane to rodzina algorytmów uczenia maszynowego, która pozwala komputerom nauczyć się mapować dane wejściowe (cechy) i wyjściowe (etykiety), podając wystarczającą liczbę przykładów.
- * Uczenie maszynowe może rozwiązać wiele trudnych problemów, w przypadku których konwencjonalna inżynieria oprogramowania zawodzi, ponieważ opiera się na danych historycznych, a nie na zrozumieniu matematycznym.
- * Nawet jeśli modele oparte na uczeniu maszynowym nigdy nie mogą być w 100% dokładne, metryki numeryczne (takie jak dokładność) mogą pomóc w śledzeniu ich wydajności.