

Wprowadzenie do sztucznej inteligencji

Dane

Rzeczywiste, dyskretne, statyczne i dynamiczne rzeczy oraz surowe obserwacje danego obszaru zainteresowania są znane jako dane. Informacje można generować po systematycznym przetwarzaniu takich danych. Dane są często identyfikowane jako wartości liczbowe w środowisku. Dane można również obserwować jako transakcyjne, fizyczne zapisy działań przedsiębiorstwa, które są uważane za podstawowy element składowy każdego systemu informatycznego. Przed użyciem ich wymagamy przetworzenia. Dane można zdefiniować jako:

Dane to symbole reprezentujące właściwości obiektów, zdarzeń i ich środowisk. Są produktami obserwacji. Obserwować to wyczuwać.

Technologia oprzyrządowania wykrywającego jest oczywiście wysoce rozwinięta. Dane to rzeczy przekazane analitykowi, badaczowi lub rozwiązującemu problem; mogą to być liczby, słowa, zdania, zapisy i założenia - tylko wszystko podane, bez względu na formę i pochodzenie. Kiedyś było to dobrze znane uczonym w większości dziedzin: niektórzy chcieli, aby dane słowo odnosiło się do faktów, zwłaszcza do odczytów instrumentów. Inni, którzy zajmują się hipotezą, dla nich dane są założeniami. Chociaż dane są na coś dowodem, nie zawsze muszą być prawdziwe; jednak trudno jest „wiedzieć”, że dane są prawdziwe lub nie. Prowadzi to do dalszego przetwarzania w celu generowania informacji i wiedzy z dostępnych danych. Na przykład temperatura o określonej godzinie w danym dniu jest pojedynczym atomem danych i traktowana jako szczególny fakt. Może istnieć kilka takich atomów, które można łączyć na różne sposoby przy użyciu standardowych operacji logiki. Istnieją jednak również uniwersalne stwierdzenia, takie jak „Każdego dnia maksymalna temperatura przekracza 30 stopni”. Jednak z logicznego punktu widzenia takie uniwersalne stwierdzenia są silniejsze niż atomy lub związki atomów, dlatego trudniej jest zapewnić ich prawdziwość. Takie dane są również wymagane do dalszego filtrowania w celu wygenerowania niezbędnych prawdziwych informacji. Przed wszystkim dane mogą być danymi empirycznymi. Bardzo trudno jest przypisać wartość prawdy fikcyjnym nieempirycznym danym.

Informacja

Kiedy dane są przetwarzane, organizowane, ustrukturyzowane lub prezentowane w określonym kontekście, tak aby były przydatne, nazywane są informacją. Chociaż istnieją informacje, które nie są danymi. Takie wyróżnione informacje można uznać za dane przetwarzane, co ułatwia podejmowanie decyzji. Przetwarzanie obejmuje agregację danych, obliczenia danych, korekty danych itp. W taki sposób, że generuje przepływ komunikatów. Informacje zwykle mają pewne znaczenie i cel. Oznacza to, że dane w kontekście można uznać za informację. Można dodać wartość do danych na kilka sposobów:

- * Kontekstualizowane: określa cel, dla którego dane zostały zebrane
- * Skategoryzowane: informuje nas o jednostkach analizy lub kluczowych składnikach danych
- * Obliczone: informuje nas, czy dane zostały przeanalizowane matematycznie czy statystycznie
- * Poprawione: informuje nas, czy błędy zostały usunięte z danych
- * Skondensowane: informuje nas, czy dane zostały streszczone w bardziej zwartej formie

Ponadto informacje mogą być przetwarzane, dostępne, generowane, przesyłane, przechowywane, wysyłane, dystrybuowane, wytwarzane i konsumowane, wyszukiwane, wykorzystywane, kompresowane i powielane. Informacje mogą być również różnego rodzaju z różnymi atrybutami. Mogą to być informacje wrażliwe, jakościowe lub ilościowe.

Wiedza, umiejętności

Wiedza jest uważana za ludzkie zrozumienie przedmiotu, który został uzyskany poprzez odpowiednie studia i doświadczenie. Informacje i dane mogą być powiązane z grupą ludzi i uważane za masę zbiorową, podczas gdy wiedza zwykle opiera się na uczeniu się, myśleniu i właściwym zrozumieniu problemu przez jednostkę. Wiedza jest uzyskiwana z informacji w sposób równoważny z danymi. Można to uznać za syntezę i integrację ludzkich procesów percepcyjnych, które pomagają im wyciągać znaczące wnioski. Wiedza jest „uzasadnioną prawdziwą wiarą” związaną z ludzkimi działaniami i jest tworzona z przepływu wiadomości. Wiedza jest zasadniczo osobista, subiektywna i z natury lokalna - znajduje się „w głowach pracowników”, a nie istnieje obiektywnie. Co więcej, wiedza może być posiadana poza ludzkim umysłem i sugeruje, że agenci są w stanie manipulować przekonaniem i osądami. Wiedza opisuje jako „prawdy i przekonania, perspektywy i koncepcje, osądy i oczekiwania, metodologie i know-how i jest w posiadaniu ludzi lub innych czynników”. Informacje to dane, które mówią o swojej działalności i o tym, jak funkcjonuje. Dodatkowy krok dotyczy informacji, aby przekształcić je w wiedzę, identyfikując trzy „I” w branży w następujący sposób:

- * Silny Wpływ (Impact): Wpływ firmy na docelową grupę użytkowników i rynek
- * Interakcje: Jak system biznesowy współdziała z użytkownikami i innymi systemami w środowisku
- * Wpływ (Influenced): wpływ na konkurencję i trendy rynkowe.

W dziedzinie zarządzania wiedzą istnieją dwa dość różne i powszechnie akceptowane rodzaje wiedzy: ukryta i jawna. Ukryta wiedza zidentyfikowana przez Polanyi to wiedza trudna do zakodowania i przekazania. Jest efemeryczna i przejściowa i „nie można jej rozdzielić na informacje ani wyszczególnić w sposób charakterystyczny dla informacji”. Ponadto wiedza ukryta jest osobista, kontekstowa i trudna do sformalizowania. Z drugiej strony, wiedza jawna jest dokładnie taką wiedzą, którą można zakodować i którą można przekazać w języku. Jest to wyraźna wiedza, że większość obecnych praktyk zarządzania wiedzą stara się i rzeczywiście może przechwytywać, pozyskiwać, tworzyć, wykorzystywać, utrzymywać, kodyfikować, przechowywać, przekazywać i udostępniać. Dane i informacje są bardzo ważnymi aspektami wiedzy. Wymaga odpowiedniego przetwarzania, aby wygenerować ustrukturyzowane znaczące informacje, aby pomóc w podejmowaniu decyzji i zdobyć wiedzę specjalistyczną w zakresie rozwiązywania problemów. Oznacza to, że poziom przetwarzania sprawia, że treść ma sens i ma zastosowanie. Poprzez odpowiednie przetwarzanie możemy generować raporty, które wspomagają podejmowanie decyzji, koncepcje uczenia się i modele rozwiązywania problemów.

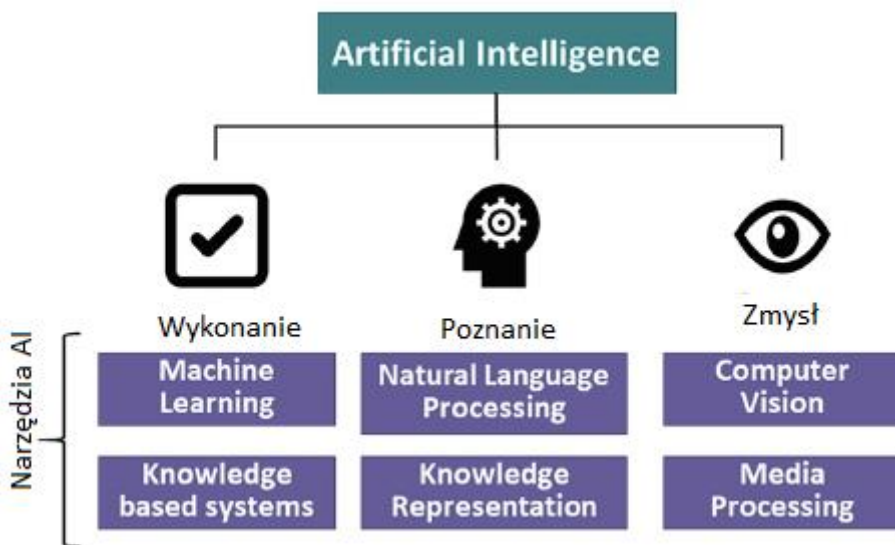
Inteligencja

Znajomość pojęć i modeli prowadzi do wyższego poziomu wiedzy zwanego mądrością. Trzeba stosować moralność, zasady i wiedzę, aby zdobywać i wykorzystywać mądrość. To wymaga czasu i wymaga pewnego rodzaju dojrzałości związanej z wiekiem i doświadczeniem. Pojęcie mądrości zostało przeszukane przez starożytnych greckich filozofów, takich jak Platon i Arystoteles; chociaż ostatnio nie był to popularny temat dyskusji. Wydaje się, że mądrość ma kilka różnych aspektów. Osoba może mieć encyklopedyczną wiedzę na temat faktów i liczb dotyczących krajów świata; ale sama ta wiedza nie uczyni tej osoby mądrą. Zamiast tego człowiek staje się mądry, stosując wiedzę do złożonych problemów o charakterze etycznym i praktycznym oraz szukając potencjalnych rozwiązań. Dalszym

ulepszeniem mądrości jest inteligencja. Inteligencja jest celem bytu, aby stać się pełnym i kompletnym sztucznie inteligentnym.

Podstawowe pojęcia sztucznej inteligencji

Sztuczna inteligencja (AI) istnieje od lat; jednak kwestia tego, gdzie można ją rozwinąć, jest kwestią dyskusji. Wraz z rozwojem technologii istnieje obecnie ogromne zapotrzebowanie na wszechstronne uczenie się człowieka w aspektach obliczeniowych - zdolne do zmiany własnego przekonania behawioralnego. Mając zdolność do decydowania, uczenia się i wpajania się w oparciu o poprzednie wydarzenia i działając na nim bardzo sumiennie. AI odnosi się do różnorodnych narzędzi i technologii, które można łączyć na różne sposoby w celu wykrywania, poznawania i działania z możliwością uczenia się na podstawie doświadczenia i dostosowywania się w czasie, jak pokazano na rysunku



Ogólnie rzecz biorąc, inteligencja jest zdolnością do zrozumienia obiektywnego świata i zastosowania wiedzy do rozwiązywania problemów. Inteligencja jednostki składa się z szerokiej gamy możliwości, takich jak: zdolność postrzegania i rozumienia obiektywnych rzeczy, obiektywnego świata i siebie; zdolność do zdobywania doświadczenia i wiedzy poprzez naukę; umiejętność zrozumienia wiedzy i zastosowania wiedzy i doświadczenia do analizy i rozwiązywania problemów; możliwości stowarzyszenia, rozumowania, oceny i podejmowania decyzji; zdolność abstrakcji językowej i uogólnienia; możliwości odkrywania, wynalazków, kreatywności i innowacji; zdolność do odpowiedniego, szybkiego i rozsądnego radzenia sobie ze złożonymi środowiskami; oraz zdolność do przewidywania i wglądu w rozwój i zmiany rzeczy. Sztuczna inteligencja nie jest nową koncepcją - w rzeczywistości większość jej podstaw teoretycznych i technologicznych była rozwijana w ciągu ostatnich 62 lat. Dla przypomnienia, za oficjalny start AI jest uważana „konferencja Dartmouth” z 1956 r. I do pewnego stopnia test Turinga wyprzedza nawet to i dostarczył myśli, jak rozpoznać „inteligentną maszynę”. Jednak podróż AI była dość burzliwa. Patrząc wstecz, nastąpił znaczny postęp w prawie wszystkich obszarach, które były przede wszystkim uważane za część sztucznej inteligencji. Spójrzmy na niektóre stymulujące rozwój pod względem praktycznego znaczenia. Systemy oparte na wiedzy były prawdopodobnie najsukcesowniejszą praktyczną gałęzią sztucznej inteligencji. W organizacjach na całym świecie wdrożono kilka aplikacji. Opracowano setki narzędzi, powszechnie nazywanych specjalistycznymi powłokami systemowymi. Takie systemy osiągnęły wystarczającą wielkość, aby stać się niezależną dyscypliną, do tego stopnia, że miały osobne kursy akademickie. Oprócz praktycznych sukcesów dziedzina ta przyczyniła się również do wzrostu samej AI. Pojęcie reprezentacji wiedzy opartej na

regułach, nacisk na rozumowanie z niepewnością, kwestie weryfikacji wiedzy dziedzinowej, uczenie maszynowe w zakresie automatycznego pozyskiwania wiedzy itp. były niektórymi obszarami rozwoju akademickiego. Kolejnym obszarem postępu było przetwarzanie języka naturalnego. Rozsądne systemy tłumaczeń są obecnie dostępne do użytku w ograniczonym kontekście, szczególnie skuteczne, jeśli do systemu można dostarczyć trochę ludzkich wskazówek. Systran jest odpowiednim przykładem, który dostarcza wewnętrzne rozwiązania językowe w czasie rzeczywistym, współpraca, wyszukiwanie, eDiscovery, zarządzanie treścią, obsługa klienta online i e-handel. Dziedzina ta przyczyniła się również do rozwoju obszaru wyszukiwania informacji. Internet jest jednym z głównych powodów zainteresowania tym obszarem, a dostępne informacje znacznie przekraczają granice ludzkiej wyobraźni. Bez zautomatyzowanej analizy i filtrowania identyfikacja i wyszukiwanie interesujących przedmiotów z tej ogromnej kopalni jest trudnym zadaniem. Sieć semantyczna, analiza treści i linków stron internetowych, eksploracja tekstu, ekstrakcja określonych informacji z dokumentów, automatyczne klasyfikowanie i spersonalizowane agenci szukające informacji zainteresowanie konkretną osobą to niektóre z aktywnych obszarów dzisiaj. Przetwarzanie mowy wygenerowało już cenne funkcjonalnie narzędzia. Obecnie dostępne są narzędzia programowe, które mogą konwertować tekst poken na tekst przetwarzany maszynowo, taki jak dokument Word. Wymagają one pewnego szkolenia i nie są jeszcze bardzo skuteczne w dostosowywaniu się do wielu głośników. Takie narzędzia są przydatne dla osób, które nie mają dobrej prędkości pisania, a co ważniejsze, dla osób niepełnosprawnych do interakcji z komputerami. Robotyka znajduje się również na ścieżce o dużym pędzie. Istnieje znacząca japońska inicjatywa, której celem jest opracowanie robotów humanoidalnych, aby pomóc osobom starszym w ich rutynowej pracy. Ten rodzaj inicjatywy przyspiesza obecnie robotykę w Japonii i USA. Honda i Sony z Japonii zbudowały roboty, które potrafią chodzić, machać, wykonywać podstawowe czynności taneczne itp. Zwierzęta robotyczne osiągnęły status komercyjny, a kilka firm sprzedaje wyrefinowane psy domowe. To, co zanotowaliśmy, to tylko część sukcesów sztucznej inteligencji. Od skromnego początku około pół wieku temu AI wzrosła w wielu wymiarach. Podczas gdy niektórzy praktycy AI dążą do osiągnięcia pierwotnego celu, jakim jest inteligencja maszyn, większość badań nad AI koncentruje się obecnie na rozwiązywaniu złożonych problemów praktycznych. Chociaż sztuczna inteligencja jest częścią naszego codziennego życia, technologia ta znajduje się w punkcie zwrotnym, głównie ze względu na ostatnie kluczowe postępy w aplikacjach do głębokiego uczenia się.

Głębokie uczenie się wykorzystuje sieci, które są zdolne do uczenia się bez nadzoru z danych, które są nieustrukturyzowane lub nieznakowane. Sieci neuronowe, które leżą u podstaw możliwości głębokiego uczenia się, stają się coraz bardziej wydajne i dokładne ze względu na dwa znaczące postępy technologiczne: bezprecedensowy dostęp do dużych zbiorów danych i wzrost mocy obliczeniowej. Skuteczność sieci neuronowych koreluje z ilością dostępnych danych. Uczenie maszynowe (ML), jeden z najbardziej ekscytujących obszarów sztucznej inteligencji, polega na opracowaniu metod obliczeniowych służących do automatycznego rozpoznawania danych - technologia ta pozwala uzyskać wgląd w to, że uczenie się jest procesem dynamicznym i możliwym dzięki przykładom i doświadczeniom w przeciwieństwie do uprzednio zdefiniowanych zasad. Podobnie jak człowiek, maszyna może zatrzymać informacje i z czasem staje się mądrzejsza. W przeciwieństwie do człowieka, maszyna nie jest skłonna do pozbawiania snu, rozpraszania uwagi, przeciążania informacji i krótkotrwałej utraty pamięci - to właśnie tam ta wpływowa technologia staje się ekscytująca. Dzięki aplikacjom w niemal każdej branży sztuczna inteligencja obiecuje znaczącą transformację istniejących modeli biznesowych przy jednoczesnym tworzeniu nowych. Na przykład w usługach finansowych wyraźne korzyści płyną z lepszej dokładności i szybkości w systemach wykrywania oszustw zoptymalizowanych pod kątem sztucznej inteligencji, które w 2020 r. Mają być rynkiem 3 mld USD. Kluczową przewagą sztucznej inteligencji nad ludzką inteligencją jest jej skalowalność, długowieczność

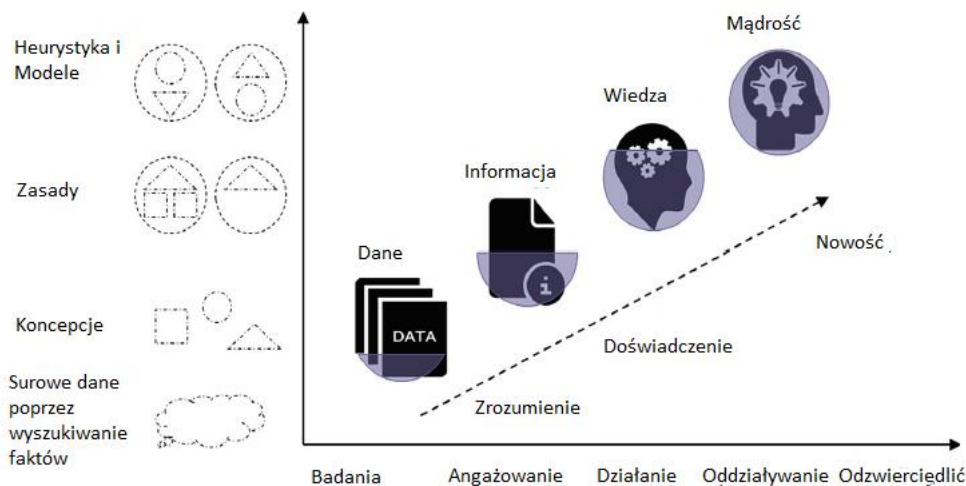
i ciągłość możliwości poprawy. Oczekuje się, że takie atrybuty znacznie zwiększą wydajność, obniżą koszty i zmniejszą liczbę błędów ludzkich. Chociaż na obiecującym etapie technologia AI ma wprowadzić nowy standard wydajności korporacyjnej, przewagi konkurencyjnej, a ostatecznie wzrostu gospodarczego. Ogromna ilość danych zgromadzonych w obecnych bazach danych ma bardzo ograniczone zastosowanie, jeśli stosowane są tylko zwykłe mechanizmy wyszukiwania. Zadawanie właściwych pytań i łączenie danych w sposób dopasowany do pytań może dostarczyć istotnych informacji. Obecnie dostępny jest cały zestaw metod, na przykład hurtownie danych, klasyczne metody statystyczne, sieci neuronowe i algorytmy uczenia maszynowego. Sztuczna inteligencja wniosła znaczący wkład w technologię danych, która jest wykorzystywana komercyjnie od kilku lat i jest często nazywana eksploracją danych.

Korzyści z AI

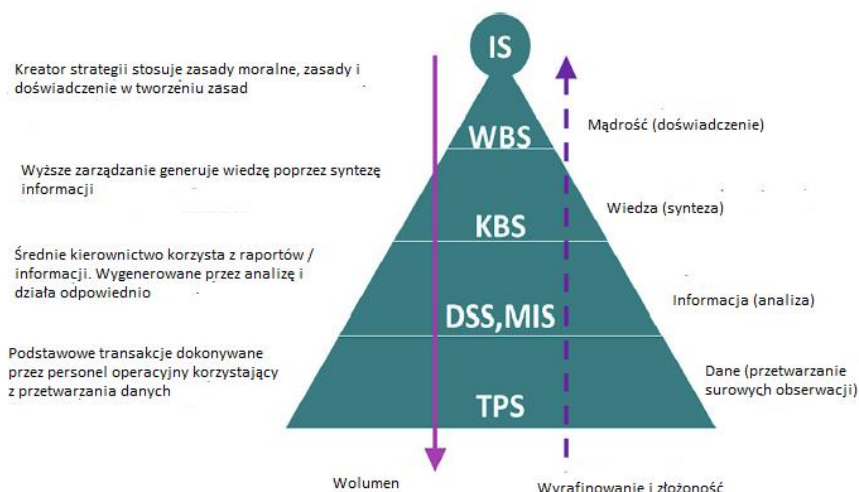
- * Odkryj znaczące i przydatne wzorce w dużych ilościach danych dowolnego rodzaju, w tym tekstu, obrazów, filmów i innych nieustrukturyzowanych danych.
- * Modele samouczące się pozwalają szybko dostosować się do zmian we wzorcach danych i podstawowych warunkach biznesowych.
- * Podejmuj lepsze decyzje szybciej, maksymalizując wartość wszystkich swoich danych i przechodząc od analizy predykcyjnej do analizy nakazowej.
- * Rozpoznawaj unikalne spostrzeżenia na temat archetypów w danych, które umożliwiają optymalną segmentację klientów i leczenie.
- * Ułatwia użytkownikom biznesowym i naukowcom danych interpretowanie danych, reagowanie na nie i korzystanie z nich dzięki lepszej wizualizacji i przejrzystości.
- * Oferuje nowe modele biznesowe i tworzenie wartości poprzez przyspieszenie innowacji poprzez odkrywanie nowych wzorców danych i pełne wykorzystanie zasobów wiedzy.

Piramida danych

Systemy AI wykorzystują techniki AI, dzięki którym osiągają kompetencje na poziomie eksperckim w rozwiązywaniu problemów w danych obszarach. Takie systemy, które wykorzystują wiedzę jednego lub więcej ekspertów do rozwiązywania problemów w określonej dziedzinie, nazywane są systemami opartymi na wiedzy lub ekspertami. Tradycyjne systemy informacyjne działają na danych i / lub informacjach. Rysunki 2 i 3 przedstawiają piramidę danych określającą relacje między danymi, informacjami, wiedzą i inteligencją. Rysunek 2 przedstawia konwergencję danych do wiedzy poprzez zastosowanie działań takich jak badanie, zaangażowanie, działanie, interakcja i refleksja.



W tym procesie człowiek zwykle zyskuje zrozumienie i doświadczenie i może wymyślać innowacyjne pomysły. Działania te są pokazane na osi X. Oś Y przedstawia formy zbieżności, którymi są surowa obserwacja, koncepcje, reguły i modele oraz heurystyka. Rysunek 3 pokazuje piramidę danych z perspektywy zarządzania.



pracownicy na poziomie operacyjnym zazwyczaj pracują w środowisku ustrukturyzowanym i wykorzystują predefiniowane procedury do przeprowadzania rutynowych transakcji biznesowych, które są podstawowymi operacjami biznesowymi. W celu przeprowadzenia rutynowej transakcji przedsiębiorstwa personel operacyjny korzysta z systemu podobnego do systemu przetwarzania transakcji (TPS). Mając całkowicie ustrukturyzowane środowisko i zestaw predefiniowanych procedur, rozwój i automatyzacja takich systemów (TPS) staje się łatwa. TPS bierze pod uwagę surowe obserwacje pola i przetwarza je w celu wygenerowania znaczących informacji. To jest poziom danych piramidy. Informacje generowane w wyniku transakcji biznesowych są analizowane w celu tworzenia rutynowych i wyjątkowych raportów, które są pomocne dla menedżerów i kadry kierowniczej w podejmowaniu decyzji. System, który to robi, nazywa się systemem informacji zarządczej (MIS). TPS i MIS działają w środowisku ustrukturyzowanym pracującym z danymi i / lub informacjami. Kierownictwo musi również podjąć decyzję, biorąc pod uwagę stosunek kosztów do korzyści różnych dostępnych alternatywnych rozwiązań, aby skutecznie wykorzystać ograniczone zasoby i ograniczenia

środowiskowe. Kategoria systemu przeznaczona do tego celu to system wspomaganie decyzji (DSS). W przeciwieństwie do TPS, który korzysta tylko z baz danych i działa w środowisku ustrukturyzowanym, DSS zwykle działa w środowisku ustrukturyzowanym do częściowo ustrukturyzowanego i wykorzystuje bazę modelu i bazę danych do optymalnego wykorzystania zasobów. Systemy takie jak TPS, MIS i DSS przeprowadzają rutynowe transakcje firmy, zapewniają szczegółową analizę generowanych informacji i wspierają proces decyzyjny firmy. Jednak systemy te nie biorą samych decyzji, ani nie uzasadniają ich odpowiednimi wyjaśnieniami i uzasadnieniami, ponieważ nie posiadają wiedzy. Kierownictwo wyższego poziomu potrzebuje wiedzy i mądrości do tworzenia polityk i strategii, a zatem istnieje zapotrzebowanie na systemy oparte na wiedzy i oparte na mądrości (KBS i WBS). Stosując etykę, zasady i osądy do podjętej decyzji i po osiągnięciu określonego poziomu dojrzałości (doświadczenie), informacje można uogólniać i przekształcać w wiedzę.

Własność autonomii

W dziedzinie informatyki systemy rozproszone są rozwijane od kilku lat i okazały się autorytatywnym środkiem do zwiększenia wydajności rozwiązywania problemów, ale także do otwarcia nowych domen aplikacji dla technologii komputerowej. Wkład AI w dziedzinie systemów rozproszonych polegał nie tylko na opracowaniu nowych algorytmów dla takich systemów, ale także na wyposażeniu komponentów systemów rozproszonych w pewien stopień autonomii. Do ostatniej dekady systemy komputerowe były postrzegane głównie jako maszyny komputerowe i magazyny danych masowych i wykorzystywane do wspomaganie decyzji. Ludzie wahali się, czy zaakceptować decyzje podejmowane przez systemy komputerowe. Jednak w niektórych domenach technicznych komputery już kontrolują duże instalacje, które w pewnym sensie obejmują podejmowanie decyzji, chociaż ludzie nadal działają jako nadzorcy. W ten sposób systemy komputerowe zaczęły już być autonomiczne. W systemach wieloagentowych, które są wkładem AI w rozproszenie, kluczową kwestią jest autonomia. Element systemu może być postrzegany lub modelowany jako agent tylko wtedy, gdy ma pewien stopień autonomii, w przeciwnym razie jest uważany za element pasywny. Autonomię można scharakteryzować za pomocą kilku funkcji. Oznacza to możliwość wyboru określonej akcji z zestawu możliwych akcji w określonej sytuacji, w tym decyzji między pozostawianiem nieaktywnym lub aktywowaniem się, gdy jest to wymagane. Ta umiejętność nazywa się proaktywnością. Kolejną ważną cechą jest zdolność do utrzymywania celów. Wymagane jest nie tylko, aby agent mógł mieć cele, które pokierują jego planowaniem i działaniami, ale także aby mógł je zmieniać, upuszczać lub przyjmować nowe. Trzecią cechą jest umiejętność aktywnej komunikacji i współpracy. Ważne jest tutaj słowo „aktywne”, ponieważ wymiana komunikatów w sensie wywołań funkcjonalnych między komponentami pasywnymi jest zwykle nazywana również komunikacją, ale nie ma to na myśli tutaj. Komunikacja i współpraca między agentami jest również zachowaniem proaktywnym; agent może rozpocząć komunikację lub szukać współpracy z innymi, gdy uzna to za konieczne. Właściwość autonomii tworzy nową relację między agentami - niezależnie od tego, czy są oni realizowani jako roboty, czy jako systemy oprogramowania - i ludźmi, którzy mają jakość partnerstwa. Agentów autonomicznych nie można już traktować jak zwykłe maszyny uruchamiane w razie potrzeby, wykonujące swoją pracę i zatrzymywane. Można postrzegać ich jako sługi, ale sługi mają własną wolę i złożone zdolności poznawcze, które również należy przyznać agentom. Musimy spodziewać się w przyszłości życia w znacznie bardziej złożonym społeczeństwie niż dzisiaj razem z agentami jako rodzajem „żywych” istot. Będą obecni w naszym codziennym życiu, na przykład jako osobisty asystent w naszym systemie komputerowym do noszenia, jako kierowcy naszego samochodu lub jako kierownicy naszego domu. Badacze AI powinni wraz z socjologami zająć się problemami, które mogą powodować ta perspektywa.

Świadomość sytuacji

W ogólnym sensie wszystkie systemy komputerowe są usytuowane, ale tradycyjne systemy istnieją w bardzo ograniczonych, dobrze określonych sytuacjach całkowicie zdeterminowanych przez człowieka. Lokalizacja stała się ważnym tematem badań dopiero wraz z pojawieniem się agentów w systemach wieloagentowych i elastycznych robotów mobilnych. Oczywiście położenie jest ściśle związane z kwestią autonomii. Tylko systemy autonomiczne muszą się lokalizować i orientować oraz działać w sytuacjach. Sytuacja tutaj oznacza zestaw wpływów ze środowiska, które są przynajmniej częściowo nieprzewidywalne, ale mają ogromne znaczenie dla systemu, tak aby musiał zareagować w odpowiedni sposób. Umieszczony system musi rozwiązać dwa główne problemy, a mianowicie, jak wyczuć sytuację i jak wybrać odpowiednią reakcję. Wykrywanie sytuacji może być dość prostym zadaniem dla zwykłego agenta oprogramowania w dobrze zdefiniowanym środowisku. Jeśli jednak myślimy o agentach działających na przykład w Internecie, sprawy stają się znacznie bardziej skomplikowane, ponieważ środowisko to jest wysoce nieustrukturyzowane i dynamiczne. Jeszcze bardziej skomplikowane jest wyczuwanie za pomocą robotów. Tutaj pierwsze sygnały fizyczne mają zostać przekształcone w dane, zadanie, które mogą powierzyć fizykom i inżynierom. Ale w następnym kroku sygnały z różnych źródeł muszą zostać zebrane razem, aby uzyskać opis sytuacji, który umożliwi odpowiedniemu rozmieszczeniu systemu odpowiednią reakcję. Zadanie to nazywa się łączeniem czujników i właśnie tam pojawiają się metody sztucznej inteligencji. Z kognitywistyki dowiadujemy się, że zrozumienie sytuacji jest bardzo złożonym procesem, który wymaga dużej wiedzy podstawowej i dużej aktywności nerwowej. Mózg konstruuje sytuację na podstawie różnych danych wejściowych. Wciąż niewiele wiemy o tym procesie i z oczywistych powodów trudno go odtworzyć. Symulacje za pomocą metod AI mogą pomóc w uzyskaniu wglądu w to. Jeśli symulacja jest możliwa, może być również wykorzystana przez sztucznie umieszczone systemy do opracowania opisu sytuacji na własne potrzeby. W przypadku systemów położonych głównym celem konstruowania opisów sytuacji jest wykorzystywanie ich do własnych działań. Podstawowe zadania samo-lokalizacji i orientacji oraz wynikające z nich zadania działania można wykonać tylko w zależności od sytuacji. Z drugiej strony oznacza to, że konstrukcja opisu sytuacji ma zawsze na celu wsparcie systemu położonego w wypełnianiu jego zadań, a nie koniec. Kryteria takie jak kompletność lub spójność nie mają pierwszeństwa, a raczej opis zaspokaja potrzeby usytuowanego systemu, jeśli pomaga on wybrać odpowiednie działania. Wymagane są rozmieszczone bazy wiedzy, składające się z szerokiej wiedzy podstawowej i fragmentów odpowiedniej wiedzy, które nie muszą być spójne ze sobą ani z wiedzą podstawową. Wiedza w tych fragmentach może być nawet reprezentowana w różnych formach i różnych ziarnistościach. Należy opracować metody wyboru odpowiedniej części wiedzy, łączenia ich oraz transformowania wiedzy z jednej formy do drugiej. Oczywiście zlokalizowane systemy muszą mieć możliwości planowania w celu wybrania sekwencji działań. Muszą także mieć zdolność uczenia się, ponieważ, jak wspomniano powyżej, wpływy, które stanowią sytuację, są częściowo nieprzewidywalne. Jednak nowe sytuacje nie różnią się od siebie całkowicie, raczej we wszystkich rozsądnych środowiskach występują między nimi podobieństwa, takie jak uczenie się, czyli wykrywanie i klasyfikowanie podobnych przypadków, ma sens. Uczenie się może poprawić zachowanie zlokalizowanego systemu. Tworzenie informacji, autonomia i lokalizacja mogą być uważane za główne obszary badań i rozwoju AI w przyszłości. Aby sprostać tym wyzwaniom, wiele pojedynczych metod musi zostać zintegrowanych z większymi systemami. Tak więc ogólny kierunek badań i rozwoju AI można scharakteryzować poprzez opracowanie złożonych systemów, które integrują różne metody i spełniają trzy wymagania.

Innowacje biznesowe z Big Data i sztuczną inteligencją

Zapotrzebowanie na dane rośnie w ciągu ostatnich kilku lat. Firmy spieszą się z przyjęciem wewnętrznych hurtowni danych i oprogramowania do analizy biznesowej i sięgają po publiczne i prywatne bazy danych w poszukiwaniu danych, aby pobudzić swoje strategie sztucznej inteligencji. Ze względu na rosnący popyt dane stają się cennym towarem, a firmy zaczynają konkurować o najbardziej

lukratywne rezerwy. Do niedawna firmy nie zdawały sobie sprawy, że siedzą w skarbcu danych i nie wiedziały, co z tym zrobić. Dzięki innowacyjnym postępom w eksploracji danych i sztucznej inteligencji firmy mogą teraz korzystać z danych wytwarzanych przez konsumentów i użytkowników. Na przykład Moz wykorzystał sztuczną inteligencję do przewidywania odejścia klientów za pomocą sieci neuronowej do głębokiego uczenia, która analizuje działania użytkownika i może przewidzieć zachowanie użytkowników. Ponieważ działania, które klienci zamierzają wykonać w systemie, są spowodowane kilkoma czynnikami z przeszłości, umożliwia to wydobycie cennych informacji biznesowych i zmniejszenie odejścia obecnych klientów, co ma ogromny wpływ na ogólny rozwój firmy. W ostatnim czasie kluczowymi źródłami danych dla dużych przedsiębiorstw były działania konsumentów online, takie jak wyszukiwane hasła, kliknięcia lub zakupy. Jednak, jak się okazuje, dane są obfite w naszych fizycznych środowiskach i doświadczeniach offline. Duże firmy, takie jak Amazon, wprowadziły korporacyjne strategie nadzoru w sklepach spożywczych. Nowe czujniki i siłowniki zainstalowane w sklepach mogą gromadzić dane na temat preferencji i zachowań konsumentów. Drony, asystenci AI, a nawet Internet przedmiotów (IoT) to narzędzia, które mogą zmienić każdą chwilę ludzkiego życia w cenne dane. Dane te stają się siłą napędową algorytmów ustalania cen, które reagują na zmiany popytu konsumpcyjnego. Uber zaczął używać tego modelu w swoim mechanizmie cenowym. Firmy, które stoją na krawędzi takich innowacji, będą miały najlepsze perspektywy na wydobycie wartości z zachowań konsumentów. Jedną z najbardziej obiecujących ścieżek jest analiza sentymentów, która wykorzystuje techniki NLP do zrozumienia dynamiki emocji i opinii użytkowników. Dzięki analizie nastrojów można również zidentyfikować pozytywne i negatywne recenzje swoich produktów na platformach e-commerce, takich jak Amazon. Ponadto znajomość nastrojów związanych z konkurencją może pomóc firmom ocenić własne wyniki i znaleźć sposoby ich poprawy. Jedną z zalet analizy nastrojów w zarządzaniu reputacją online jest automatyzacja, ponieważ ręczne przetworzenie wielu opinii użytkowników może być trudne. Zamieniając informacje zwrotne w dane, które zostaną przekazane do twojego oprogramowania Business Intelligence to jedno z najbardziej wydajnych rozwiązań, które wyróżni Cię na tle konkurencji. Od chatbotów i inteligentnych generatorów narracji po narzędzia analizy biznesowej - sztuczna inteligencja staje się prawdziwą przewagą konkurencyjną dla firm promujących automatyzację, redukcję kosztów i inteligentne podejmowanie decyzji. Jednak, aby opracować strategie sztucznej inteligencji i wyszkolić modele uczenia maszynowego, firmy potrzebują danych wysokiej jakości. Facebook i Google rozwiązały ten problem, wykorzystując model „użytkownik w pętli”, w którym użytkownicy generują dla nich dane za pośrednictwem postów, komentarzy lub zapytań. Niektóre firmy uzyskują dostęp do danych poprzez sięganie do publicznych i komercyjnych baz danych, crowdsourcingowe usługi gromadzenia i klasyfikacji danych, współpracę z firmami opartymi na danych itp. Niezależnie od tego, które podejście najlepiej pasuje do twojego modelu biznesowego, musisz wprowadzić skuteczne strategie akwizycji danych, aby wykorzystać moc AI.

Pokrywanie się sztucznej inteligencji z innymi dziedzinami

Sztuczna inteligencja polega na uczynieniu maszyny inteligentną i podejmowaniu decyzji z uzasadnieniem. To pole używa również danych i uczy maszyny. Sztuczna inteligencja jest dziedziną, która jest wszechobecnie stosowana w większości innych dziedzin i może przyczyniać się do dowolnej dziedziny. Ma zdolność uczenia się z dużej ilości danych, siłą symulowania zachowań inspirowanych naturą oprócz typowych inteligentnych modeli i algorytmów. Dzięki temu sztuczna inteligencja ma uniwersalne zastosowanie tam, gdzie zawodzi typowy model formalny. Być może najbardziej znaczącą różnicą są moc obliczeniowa i ilość danych, które możemy gromadzić i analizować w porównaniu z poprzednimi dekadami. Smartfon, który dziś łatwo mieści się w dłoni, może przechowywać i przetwarzać więcej danych niż komputer mainframe z lat 60. XX wieku, który zajmował kilka pomieszczeń. Zamiast polegać na dokładnie wyselekcjonowanych i małych zestawach danych, możemy

wykorzystywać duże i niezorganizowane dane z tysiącami parametrów do trenowania algorytmów i rysowania prognoz. Ilość i jakość danych są tym, co odróżnia nowoczesne techniki uczenia maszynowego od statystyki. Podczas gdy statystyki zwykle polegają na kilku zmiennych, aby uchwycić wzór, uczenie maszynowe może być skutecznie wykorzystane z tysiącami cech danych. Uczenie maszynowe jest uważane za integralny element informatyki i dziedzinę związaną z ICT. W dziedzinie uczenia maszynowego nacisk kładziony jest na różne algorytmy i techniki, dzięki którym maszyna uczy się automatycznie na podstawie danych. Później wyniki te są wykorzystywane w interpretacji i zastosowaniu danych do rozwiązywania problemów. Jednak dziedzina uczenia maszynowego stosuje pewne techniki statystyczne i matematyczne. Pojęcie data science powstało w latach 60. XX wieku. W miarę rozwoju nauki o danych i zdobywania nowych „instrumentów” z czasem główny cel biznesowy nadal koncentruje się na znajdowaniu użytecznych wzorców i uzyskiwaniu cennych informacji z danych. Dzisiaj, data science jest wykorzystywana w wielu branżach i pomocach w różnych problemach analitycznych. Na przykład w marketingu badanie wieku, płci, lokalizacji i zachowania klientów pozwala na prowadzenie ściśle ukierunkowanych kampanii, oceniając, ile klientów jest skłonnych dokonać zakupu lub odejść. W bankowości znalezienie odległych działań klienta pomaga wykryć oszustwo. W służbie zdrowia analiza dokumentacji medycznej pacjentów może wykazać prawdopodobieństwo wystąpienia chorób itp. Eksploracja danych jest również ściśle związana z uczeniem maszynowym i sztuczną inteligencją. Termin „eksploracja danych” jest niedokładnym terminem i nie brzmi tak jak to oznacza. Zamiast samej eksploracji danych dyscyplina polega na tworzeniu algorytmów w celu wydobywania cennych informacji z dużych i prawdopodobnie nieustrukturyzowanych danych. Podstawowym problemem eksploracji danych jest mapowanie dostępnych danych i przekształcanie ich w strawne wzorce. Eksploracja danych jest uważana za część szerszego procesu zwanego wiedzą w bazach danych (KDD), które zostało wprowadzone w 1984 roku przez Gregory'ego Piatetsky-Shapiro. Niektóre z typowych technik obejmują rozpoznawanie wzorców, klasyfikację, dzielenie i grupowanie wraz z kilkoma modelami statystycznymi. Oznacza to, że eksploracja danych również w pewnym stopniu pokrywa się ze statystykami. W erze zdominowanej przez media społecznościowe personalizacja klienta staje się jednym z głównych źródeł przewagi konkurencyjnej dla firm oferujących swoje produkty i usługi online. Narzędzia analizy konsumenckiej i najnowocześniejsze oprogramowanie AI do silników rekomendacji to główne czynniki zmieniające grę, które umożliwiają efektywną personalizację w biznesie. Dane dotyczące preferencji użytkownika, zainteresowań oraz zachowań w czasie rzeczywistym i przeszłości można teraz łatwo gromadzić, przechowywać i analizować za pomocą narzędzi analizy biznesowej i algorytmów AI. Na przykład informacje pochodzące z tych danych pozwalają sprzedawcom dostarczać odpowiednie treści użytkownikom witryny, projektantom gier wideo w celu dostosowania trudności i funkcji gry do graczy lub silniki rekomendacji sugerujące muzykę, filmy lub produkty, które mogą się podobać konsumentom. Personalizacja oparta na danych staje się zatem doskonałym narzędziem do zatrzymywania konsumentów i oferowania im produktów, usług i funkcji, których naprawdę szukają.

Problemy z etyką i prywatnością

Większość aplikacji AI wymaga ogromnej ilości danych, aby uczyć się i podejmować inteligentne decyzje. Sztuczna inteligencja zajmuje ważne miejsce w większości sektorów ze względu na jej potencjał w zakresie radykalnie ulepszonych usług, przełomów komercyjnych i zysków finansowych. W przyszłości napotkamy szereg dylematów prawnych i etycznych w poszukiwaniu równowagi między znacznymi postępami społecznymi w imię sztucznej inteligencji a podstawowymi prawami do prywatności. Dane i algorytmy składające się na sztuczną inteligencję nie mogą być po prostu dokładne i wydajne; muszą również zaspokoić obawy dotyczące prywatności i spełnić wymogi prawne. Kwestie związane z danymi mogą być wyraźne w ściśle regulowanych branżach, takich jak ubezpieczenia, które przechodzą od modelu historycznego opartego na łączeniu ryzyka do podejścia, które obejmuje

elementy przewidujące określone ryzyko. Ale niektóre atrybuty są wyłączone. Na przykład, podczas gdy czynniki płci i religii mogą być wykorzystane do przewidywania niektórych zagrożeń, są one nie do zaakceptowania przez organy regulacyjne w niektórych aplikacjach i jurysdykcjach. Ponieważ technologia konkuruje z oczekiwaniami i preferencjami konsumentów, firmy stawiają coraz cieńszą granicę między swoimi inicjatywami sztucznej inteligencji, ochroną prywatności i obsługą klienta. Na przykład dostawcy usług finansowych używają technologii rozpoznawania głosu do identyfikacji klientów przez telefon, aby zaoszczędzić czas na sprawdzaniu tożsamości. Klienci chętnie przyjmują to doświadczenie, a nie zawahają się, po części dlatego, że cenią usługę i ufają firmie, że nie wykorzysta możliwości lub danych, które ją umożliwiają. Nowe przepisy Unii Europejskiej dotyczące ochrony danych, które weszły w życie w maju 2018 r., wzmocnią nasze prawa do prywatności, a jednocześnie zaostrzą wymagania wobec osób przetwarzających takie dane. Organizacje będą ponosić większą odpowiedzialność za przetwarzanie danych osobowych zgodnie z rozporządzeniem, a wymogi dotyczące przejrzystości będą bardziej rygorystyczne. Wraz z intensyfikacją wymagań rośnie zapotrzebowanie na dane. Systemy oparte na sztucznej inteligencji mogą stać się inteligentne tylko wtedy, gdy mają wystarczającą ilość odpowiednich danych, aby się z nich uczyć. Inteligentny chatbot analizuje wszystkie dostarczane informacje - kombinację pytań zadawanych przez klientów i odpowiedzi przekazywanych przez obsługę klienta. Dzięki swojej analizie chatbot może „zrozumieć”, o co pyta klient, i dlatego może udzielić sensownej odpowiedzi. Im większa ilość informacji, na których chatbot może oprzeć swoją analizę, tym lepsza i dokładniejsza będzie udzielona odpowiedź. Przepisy ogólnego rozporządzenia o ochronie danych (RODO) 2 regulują obowiązki administratora danych i prawa osoby, której dane dotyczą podczas przetwarzania danych osobowych. RODO ma zatem zastosowanie, gdy sztuczna inteligencja jest opracowywana za pomocą danych osobowych, a także gdy jest wykorzystywana do analizy lub podejmowania decyzji dotyczących osób fizycznych. Zasady przetwarzania danych osobowych opierają się na niektórych podstawowych zasadach. W art. 5 RODO wymieniono zasady mające zastosowanie do całego przetwarzania danych osobowych. Istotą tych zasad jest to, że dane osobowe powinny być wykorzystywane w sposób, który chroni prywatność osoby, której dane dotyczą, w najlepszy możliwy sposób, i że każdy ma prawo decydować o sposobie wykorzystania jego danych osobowych. Wykorzystanie danych osobowych w rozwoju sztucznej inteligencji stanowi wyzwanie dla kilku z tych zasad. Podsumowując, zasady te wymagają, aby dane osobowe były:

- * Przetwarzane w sposób zgodny z prawem, uczciwy i przejrzysty (zasada legalności, sprawiedliwości i przejrzystości)
- * Zebrane do konkretnych, wyraźnie określonych i uzasadnionych celów i nie traktowane w nowy sposób niezgodny z tymi celami (zasada ograniczenia celu)
- * Odpowiednie, odpowiednie i ograniczone do tego, co jest konieczne do osiągnięcia celów, dla których jest przetwarzany (zasada minimalizacji danych)
- * Prawidłowe i w razie potrzeby zaktualizowane (zasada dokładności)
- * Nie przechowywane w możliwej do zidentyfikowania formie przez okresy dłuższe niż jest to konieczne do celów (zasada odnosząca się do okresów przechowywania danych)
- * Przetwarzane w sposób zapewniający odpowiednią ochronę danych osobowych (zasada integralności i poufności)

Sztuczna inteligencja to szybko rozwijająca się technologia. To samo dotyczy narzędzi i metod, które mogą pomóc sprostać wyzwaniom związanym z ochroną danych przy użyciu sztucznej inteligencji. Zebraliśmy kilka przykładów ilustrujących niektóre z dostępnych opcji. Metody te nie zostały ocenione

w praktyce, ale ocenione pod kątem ich potencjalnego potencjału. Oznacza to, że pod względem technicznym mogą być dzisiaj nieodpowiednie, ale koncepcje są ekscytujące i mają potencjał do dalszych badań i przyszłego wykorzystania.

AI i analityka predykcyjna

Analityka predykcyjna i sztuczna inteligencja to dwie różne rzeczy. Po połączeniu wydobywają ze sobą to, co najlepsze. AI umożliwia analitykę predykcyjną, by była szybsza, inteligentniejsza i bardziej praktyczna niż kiedykolwiek wcześniej. Gdy firmy chcą tworzyć prognozy oparte na danych o przyszłych zdarzeniach, opierają się na analizie predykcyjnej. W erze dużych zbiorów danych analiza predykcyjna szybko staje się ważną częścią wielu firm i funkcji. Analityka predykcyjna polega na wykorzystaniu danych historycznych do prognozowania. Najlepszym przykładem analizy predykcyjnej jest ocena wiarygodności kredytowej. Wynik oparty jest na historii kredytowej w przeszłości i służy do przewidywania prawdopodobieństwa spłaty zadłużenia. Chociaż analityka predykcyjna jest używana od dziesięcioleci w usługach finansowych, dopiero niedawno stała się kluczowym narzędziem w innych firmach. Postęp technologii gromadzenia i przetwarzania danych umożliwił zastosowanie analizy predykcyjnej do prawie każdego aspektu działalności, od logistyki po sprzedaż i zasoby ludzkie. Podstawą analityki predykcyjnej jest model. Chociaż techniki statystyczne użyte do stworzenia modelu zależą od konkretnego zadania, dzielą się one na dwa szerokie typy. Pierwszy to model regresji, który służy do oceny korelacji między określonymi zmiennymi a wynikami. Uzyskane współczynniki dają skwantyfikowaną miarę tego związku, w efekcie, jak prawdopodobne jest dane wynik opiera się na zestawie zmiennych. Innym typem modelu jest model klasyfikacyjny. Tam, gdzie modele regresji przypisują prawdopodobieństwo do zdarzenia, modele klasyfikacji przewidują, czy raczej należą do tej czy innej kategorii. Modelowanie predykcyjne i analizy są już dostępne. Brakowało jednak trzech rzeczy, które są ważne, aby osiągnąć prawdziwą wartość marketingową: skala, szybkość i zastosowanie. Właśnie wtedy w grę wchodzi AI. Dzięki sztucznej inteligencji modele predykcyjne mogą zapewniać niesamowitą ilość informacji w czasie rzeczywistym. Takie modele mogą uwzględniać znacznie więcej informacji niż kiedykolwiek wcześniej, co zwiększa ich wydajność precyzyjne i wykonalne. Ponadto AI może oceniać miliardy zmiennych w czasie rzeczywistym i podejmować jednocześnie decyzje o analizie ogromnej ilości możliwości marketingowych na sekundę. Bez sztucznej inteligencji model predykcyjny nie jest w stanie szybko zrozumieć ilości danych, a modele predykcyjne nie mają zdolności „poznawczych” do działania.

Obszary zastosowania

Zarządzanie relacjami z klientem (CRM). Korzystając z połączenia analizy regresji i technik klastrowania, narzędzia CRM mogą podzielić klientów firmy na kohorty na podstawie ich danych demograficznych i miejsca ich cyklu życia, umożliwiając ukierunkowanie działań marketingowych w najbardziej prawdopodobny sposób być efektywnym.

Wykrywanie wartości odstających i oszustw. Tam, gdzie najbardziej predykcyjne aplikacje analityczne szukają podstawowych wzorców, wykrywanie anomalii szuka elementów, które wystają. Usługi finansowe używają go do wykrywania oszustw od lat, ale te same techniki statystyczne są przydatne również w innych aplikacjach, w tym w badaniach medycznych i farmaceutycznych.

Przewidywanie popytu. Ważnym, ale trudnym zadaniem dla każdej firmy jest przewidywanie popytu na nowe produkty i usługi. Wcześniej tego rodzaju prognozy dokonywano przy użyciu danych szeregów czasowych do tworzenia ogólnych prognoz, ale teraz detaliści są w stanie anonimizować dane wyszukiwania, aby przewidzieć sprzedaż danego produktu na poziomie regionalnym.

Poprawa procesów. W przypadku producentów, producentów energii i innych firm, które polegają na skomplikowanych i wrażliwych maszynach, analizy predykcyjne mogą poprawić wydajność, przewidując, które maszyny i części mogą wymagać konserwacji. Wykorzystując historyczne dane dotyczące wydajności i dane z czujników w czasie rzeczywistym, te modele predykcyjne mogą poprawić wydajność i skrócić przestoje, pomagając jednocześnie uniknąć poważnych przestojów w pracy, które mogą wystąpić w przypadku nieoczekiwanej awarii głównych systemów.

Budowanie silników rekomendacji. W celu zwiększenia lojalności i zaangażowania użytkowników opiera się na spersonalizowanych zaleceniach usług przesyłania strumieniowego, sprzedawców internetowych, serwisów randkowych i innych. Wspólne techniki filtrowania wykorzystują połączenie wcześniejszych zachowań i podobieństwa do innych użytkowników w celu wygenerowania rekomendacji, podczas gdy filtrowanie oparte na treści przypisuje cechy do elementów i rekomenduje najnowsze elementy na podstawie ich podobieństwa do poprzednich elementów.

Poprawa czasu wynajmu i utrzymania. Firmy mogą wykorzystywać dane z systemów zasobów ludzkich w celu optymalizacji procesu rekrutacji i identyfikowania kandydatów, którzy mogą zostać przeoczeni przez osoby sprawdzające. Ponadto niektóre działy używają kombinacji danych dotyczących wydajności i profili osobowości, aby określić, kiedy pracownicy mogą odejść lub przewidzieć potencjalne konflikty, aby można je było proaktywnie rozwiązać.

Grupowanie lub segmentacja

Grupowanie to proces organizowania obiektów w grupy, których członkowie są w pewien sposób podobni. Natomiast segmentacja klientów to praktyka polegająca na dzieleniu bazy klientów na grupy osób, które są podobne w specyficzny sposób związany z marketingiem, takie jak wiek, płeć, zainteresowania, nawyki zakupowe i tak dalej. Segmentacja lub klastrowanie klientów jest przydatne na wiele sposobów. Można go wykorzystać do marketingu ukierunkowanego. Czasami budując model predykcyjny, raczej efektywne jest grupowanie danych i budowanie osobnego modelu predykcyjnego dla każdego klastra. Klastrowanie jest techniką nieukierunkowanej eksploracji danych. Oznacza to, że można go użyć do identyfikacji ukrytych wzorów i struktur w danych bez formułowania konkretnej hipotezy. W klastrowaniu nie ma zmiennej docelowej. Na przykład sprzedawca artykułów spożywczych nie próbował aktywnie zidentyfikować miłośników świeżej żywności na początku analizy. Po prostu starał się zrozumieć różne zachowania zakupowe jego bazy klientów. Grupowanie odbywa się w celu zidentyfikowania podobieństw w odniesieniu do precyzyjnych zachowań lub wymiarów. Na przykład chcemy zidentyfikować segmenty klientów o podobnych zachowaniach zakupowych. W związku z tym grupowanie przeprowadzono przy użyciu zmiennych reprezentujących wzorce zakupowe klientów. Analiza skupień może być wykorzystana do odkrycia struktur w danych bez podania wyjaśnienia lub interpretacji. Analiza skupień po prostu odkrywa wzorce w danych bez wyjaśnienia, dlaczego one istnieją. Powstałe klastry same w sobie nie mają znaczenia. Muszą być szeroko profilowane, aby budować swoją tożsamość, to znaczy zrozumieć, co reprezentują i czym różnią się od populacji rodziców. Klastrowanie służy przede wszystkim do segmentacji, czy to klienta, produktu czy sklepu. Na przykład produkty można grupować w hierarchiczne grupy w oparciu o ich atrybuty, takie jak zastosowanie, rozmiar, marka, smak itp. ; sklepy o podobnych cechach - sprzedaż, wielkość, baza klientów itp. - mogą być grupowane razem. Procedura klastrowania może być hierarchiczna, gdzie klastrowanie charakteryzuje się rozwojem hierarchii lub struktury treelike. Grupowanie aglomeracyjne rozpoczyna się od każdego obiektu w osobnym klastrze, a klastry są tworzone przez grupowanie obiektów w coraz większe klastry. Z drugiej strony klastrowanie dzielące zaczyna się od wszystkich obiektów zgrupowanych w jednym klastrze, a klastry są następnie dzielone lub dzielone, aż każdy obiekt znajdzie się w osobnym klastrze. Klastrowanie K-oznacza to klastrowanie niehierarchiczne i jest to procedura, która najpierw przypisuje lub wyznacza centrum klastra, a następnie grupuje wszystkie

obiekty w ramach wcześniej określonej wartości progowej razem pracując od środka. Decyzja o liczbie klastrow opiera się na względach teoretycznych lub praktycznych. W hierarchicznym klastrowaniu jako kryteria można zastosować odległości, w których klastry są łączone. W klastrowaniu niehierarchicznym stosunek całkowitej wariancji wewnątrz grupy do wariancji między grupami można wykreślić na podstawie liczby klastrow. Interpretowanie i profilowanie klastrow obejmuje badanie centroidów klastrowych. Centroidy reprezentują średnie wartości obiektów zawartych w gromadzie na każdej ze zmiennych. Centroidy można przypisać nazwą lub etykietą. Aby ocenić wiarygodność i trafność, należy przeprowadzić analizę skupień na tych samych danych, stosując różne miary odległości i porównać wyniki określić stabilność rozwiązań. Losowe dzielenie danych na połowy i wykonywanie grupowania osobno dla każdej połowy i porównywanie centroidów klastrowych w dwóch podpróbkach jest jednym z moich ulubionych sposobów. W hierarchicznym klastrowaniu rozwiązanie może zależeć od kolejności spraw w zbiorze danych. Aby osiągnąć najlepsze wyniki, wykonaj wiele przebiegów, stosując inną kolejność spraw, aż rozwiązanie się ustabilizuje. Grupowanie może być również wykorzystywane do wykrywania anomalii, na przykład w przypadku dentystycznych transakcji oszustw. Metody wykrywania klastra mogą być stosowane w próbkę zawierającej tylko prawidłowe transakcje w celu określenia kształtu i wielkości „normalnego” klastra. Kiedy pojawia się transakcja, która z jakiegokolwiek powodu wychodzi poza klaster, jest to podejrzane. To podejście zostało zastosowane w medycynie do wykrywania obecności nieprawidłowych komórek w próbkach tkanek oraz w telekomunikacji do wykrywania wzorców wywoływania wskazujących na oszustwo. Grupowanie jest często używane do dzielenia dużego zestawu danych na mniejsze grupy, które są bardziej podatne na inne techniki. Na przykład wyniki regresji logistycznej można poprawić, wykonując ją osobno dla mniejszych klastrow, które zachowują się inaczej i mogą mieć nieco inne rozkłady. Podsumowując, klastrowanie jest potężną techniką do badania struktur wzorców w danych i ma szerokie zastosowanie w analizie biznesowej. Istnieją różne metody grupowania. Analityk powinien znać wiele algorytmów klastrowania i być w stanie zastosować najbardziej odpowiednią technikę zgodnie z potrzebami biznesowymi.

Psychograficzne osobowości

Psychografia to wskaźniki zainteresowań, zachowań, postaw i opinii, które pomagają zrozumieć, dlaczego dana osoba może / nie może kupić produkt. Dane psychograficzne w połączeniu z danymi demograficznymi mogą dać prawie pełny obraz osoby i pomóc w wyborze rodzaju produktów, które mogą spodobać się tej osobie. Psychograficzne parametry targetowania dla danej osoby są definiowane przez psychologiczną tendencję grupy ludzi do zachowania się w określony sposób lub do podobnych rzeczy. Tak więc dla młodej matki parametry psychograficzne obejmowałyby skłonność do poszukiwania zasobów, które dałyby jej wiedzę na temat opieki nad dzieckiem. W świecie internetowym wskaźniki definiujące psychikę danej osoby obejmowałyby przeszłe czynności związane z przeglądaniem, aktywność na stronie internetowej, przeszłe zakupy, deklarowane zainteresowanie stronami społecznościowymi i inne tego rodzaju dane. Dane psychograficzne, zebrane w ten sposób i poskładane razem, mogą dać bardzo dobry wgląd w to, jakie produkty mogą być zainteresowane lub mogą być zakupione przez osobę. Segmentacja rynku to proces dzielenia rynku na segmenty lub grupy konsumentów, którzy są podobni, ale różni się od konsumentów w innych grupach. Segmentacja dzieli rynek na podgrupy. Marketing docelowy polega na decydowaniu, które segmenty są najbardziej rentowne. Ponadto pozycjonowanie obejmuje tworzenie wizerunku produktu, który będzie atrakcyjny dla rynku docelowego lub kilku rynków docelowych. Segmentacja psychograficzna pomaga konstruować produkty lub pozycjonować je w sposób, który czyni je bardziej atrakcyjnymi niż konkurenci. Tworzenie map percepcyjnych pomaga zrozumieć, w jaki sposób konsumenci postrzegają Twoją markę, i umożliwia pozycjonowanie marki w celu uzyskania maksymalnych korzyści. AI gromadzi klientów w pulach odbiorców na podstawie punktów kontaktowych i analizy nastrojów, które

pomagają marketerom zrozumieć, jak różne segmenty klientów mogą zareagować na post społecznościowy, billboard lub blog. Analizując sposób, w jaki klienci rozmawiają ze sobą, może sugerować frazy i nastroje, które najlepiej współgrają z każdym segmentem odbiorców.