

Data science w piekarni



Data science to inaczej nauka o danych. To dziedzina analiz, która w ostatnich latach w oszałamiającym tempie rozwinęła się na świecie. Rozwój ten, data science zawdzięcza ogromnemu rozwojowi Internetu i działającym w nim serwisom społecznościowym, takim jak LinkedIn czy Twitter. W sieci dokonuje się dużo transakcji detalicznych, działają serwisy informacyjne i ogromne agencje prowadzące kampanie reklamowe na niespotykaną dotąd skalę.

Na serwerach całego świata w każdej sekundzie dokonują się miliardy kliknięć. Każdy wybór klientów, dzięki wielkim mocom obliczeniowym, jest pieczołowicie zapisywany i analizowany.

Dla firm działających na masową skalę w Internecie, każde zachowanie klienta, każda jego prawidłowość jest skrupulatnie analizowana i zapisywana na koncie tego klienta.

Jeżeli niedawno szukaliśmy w Internecie odkurzacza, odkurzacze będą za nami „chodzyli”. Gdzie nie wejdziemy, zawsze pojawi się dyskretna reklama z odkurzaczami. Jeżeli nie znosimy celebrytów, a uwielbiamy piłkę nożną, algorytmy mediów informacyjnych będą nam podsuwały najnowsze futbolowe nowinki i izolowały nas od skandali nie lubianych celebrytów. Liczy się czas, jak długo przytrzymamy klienta na stronie, jak często będzie do nas wracał. Dostosowując się do potrzeb klientów witryny internetowe uzależniają ich od siebie zyskując przewagę konkurencyjną.

Co to ma wspólnego z piekarnią?

Przykład aktywności data science w Internecie nie jest zbyt trafiony w kontekście piekarni i cukierni. Użyłem tego przykładu, ponieważ ta aktywność jest bardzo widoczna dla wszystkich. Mimo że tego nie widać, data science działa we wszystkich dziedzinach gospodarki.

Statystyczna analiza procesów szczególnie rozwinęła się tam, gdzie dokonuje się operacji masowych, w sieciach komórkowych, dystrybucji gazu i energii elektrycznej, w bankach, w sieciach handlowych i w logistyce. Dzisiaj używanie data science jest warunkiem utrzymania się na coraz bardziej konkurencyjnym rynku.

Data science w mojej piekarni?

Zostawmy zgiełk i jazgot wielkiego świata, wróćmy do zacisza naszej piekarni. Tutaj też dokonały się zmiany w technologii produkcji i dystrybucji. Jednak cały proces sprzedaży, cykl produkcji i obrotu surowcami pozostał niezmieniony od lat. Czy w takim środowisku opłaca się wprowadzać jakieś nowinki z innego świata? Pomysł wydaje się niedorzeczny!

Czym naprawdę jest data science?

Tej dziedziny wiedzy nie wynaleziono wczoraj, istnieje od ponad 100 lat, kryjąc się pod różnymi nazwami. Siedemdziesiąt lat temu to była statystyka, w latach 60. zabłysła jako

ekonometria i optymalizacja matematyczna, by w początkach lat 80. pojawić się jako modna wtedy Six Sigma. Za każdym razem statystyczna rewolucja rozbijała się na kamieniach mozołnych obliczeń i siermiężnej systematyki. Wciąż trzeba było prowadzić prace analityczne, a na to brakowało cierpliwości i powszechnej wiedzy. Za każdym razem brakowało czegoś, żeby przedsięwzięcie analityczne zaczęło żyć własnym życiem.

Cyfrowa piekarnia?

Tym, co różni piekarnie i cukiernie z lat 20. od piekarni z roku 2019 jest to, że ta ostatnia oprócz przysmaków produkuje ogromne ilości danych cyfrowych. Kasa fiskalna generuje przy jednorazowej sprzedaży tuzin informacji, takich jak: cena, godzina transakcji, indeksy asortymentowe, nazwisko obsługującego sprzedaż. Jeżeli istnieje kilka sklepów, łącznie, dziennie dokonywanych jest tysiące transakcji. Rozejrzyjmy się dalej. Mamy setki danych z komputerów sterujących piecami gazowymi, mamy dane z systemów recepturowych, inwentaryzacji, dane z komputerów samochodowych. Wreszcie mamy informacje z inwentaryzacji zwrotów produktów. Ta ostatnia informacja mówi nam ile i jakie produkty się nie sprzedały. Czyli coś zostało źle przewidziane. Podjęliśmy złą decyzję, teraz musimy za nią zapłacić.

Ogromna ilość danych z cyfrowej piekarni zebrana tu i teraz jest odkładana na dyski komputerów, nawet bez naszej wiedzy. Po prostu tak musi być, to wymóg urzędu skarbowego, to rejestry serwisowe urządzeń czy pamięć logów programów sprzedażowych i produkcyjnych, o której nawet nie wiemy. Każdy cyfrowy obraz dnia tu i teraz łączy się z innymi wcześniejszymi dniami tworząc niekończące się szeregi czasowe danych.

Cyfrowy świat zewnętrzny oddziałuje na piekarnię: kursy walut, ceny gazu, a pogoda, zbliżające się święta i zmiana czasu na letni wpływają bezpośrednio na wielkość sprzedaży, techniczny koszt wytworzenia i zwroty produktów.

Te dane to skarbnica informacji o popycie, zachowaniach klientów, o produkcji i sprawności dystrybucji. Współzależność zjawisk jest widoczna gołym okiem, wiadomo że coś z czegoś wynika. Jednak dokładnie nie wiemy co, nie mamy czasu na analizę, na statystykę. Cierpimy z uwagi na naszą niewiedzę.

Sytuacja przypomina zdarzenie, kiedy mały jacht wycieczkowy rozbił się na tropikalnej, bezludnej wyspie. Po dwóch miesiącach większość rozbitków zmarła z głodu i wycieńczenia. Na wyspie była bo-

gata roślinność i słodka woda, a w morzu pływały ryby. Nikt z robitków nie potrafił skorzystać z tych zasobów.

Cukiernik statystykiem?

Tym, co różni piekarnie z roku 1920 od piekarni z 2019 roku jest technologia obróbki danych. Kiedyś piekarz prowadził księgi handlowe, zapisywał otówkiem operacje i prowadził na marginesie obliczenia. Problem polegał na tym, że za każdym razem nasz piekarz musiał wszystkie obliczenia prowadzić od nowa.

Dzisiaj w cyfrowej rewolucji nie chodzi o to, że mając gotowe formuły w Excelu podstawiamy dane i uzyskujemy wynik. Dzisiejsza technologia dostarcza nam informacji – to wszystko, nic nie musimy robić, nic nie musimy wiedzieć.

To tak jakby piekarz był zdziwiony, że za każdym razem, gdy wsiada do samochodu dostawczego, nie musi od nowa składać silnika tego samochodu. Technologia ma dostarczać informacje, a samochód ma jeździć. Nie musimy znać się na wszystkim.

Są firmy, które dałyby wszystko, aby wiedzieć jak dużo surowców muszą zamówić do produkcji. Są banki, które zrobiłyby wiele, żeby wiedzieć, jak zachowają się klienci w odpowiedzi na podwyżkę opłat. Wreszcie są cukiernie, które chciałyby wiedzieć, ile należy wyprodukować, żeby potem, wieczorem nie trzeba było tych produktów utylizować. Informacja o przyszłości zawsze była ważna. Dzisiaj informacja o przyszłości staje się dostępna w zasięgu ręki.

Kim jest Badacz Danych?

Data Scientist – inaczej Badacz Danych – to rola, która dzisiaj w rankingu zawodów daje najwyższy poziom zarobków. Jest wiele rodzajów badaczy danych. Część z nich ma stałe zatrudnienie, są to najczęściej tropiciele defraudacji bankowych, optymalizatorzy i kontrolerzy procesów. Większość Data Scientist to jednak wędrowcy, tak zwani: „cyfrowi nomadzi”, uczestniczą w krótkich, dobrze płatnych projektach technologicznych.

W ich przypadku przedmiotem takich projektów jest tworzenie samouczących się modeli prognostycznych lub inteligentnych, autonomicznych systemów optymalizacyjnych.

Odchodzą pozostawiając po sobie w pełni autonomiczne, bezobsługowe systemy uczące się, które sprawnie i za darmo dostarczają kluczowych informacji o przyszłości. Czas życia modeli, które tworzą, jest nieograniczony i zależy wyłącznie od charakterystyki procesu, w którym żyją. Dopóki nie zmieni się technologia produkcji, organizacja dystrybucji lub sprzedaży, algorytmy modeli będą funkcjonowały wiecznie.

W następnych odcinkach tego cyklu przedstawię kilka praktycznych zastosowań metod data science w piekarni i cukierni. Pokażę również jak samemu zbudować proste modele prognostyczne w Excelu.

Wojciech Moszczyński



1994-2019
KOMPLET Polska

Z dobrego wypieczesz najlepsze!

Tłusty Czwartek

KOMPLET
Bellkule

100% mieszanka do produkcji delikatnych małych pączków.

KOMPLET
Boleros

100% mieszanka do produkcji delikatnych kulek serowych.





KOMPLET Polska

www.komplet.pl