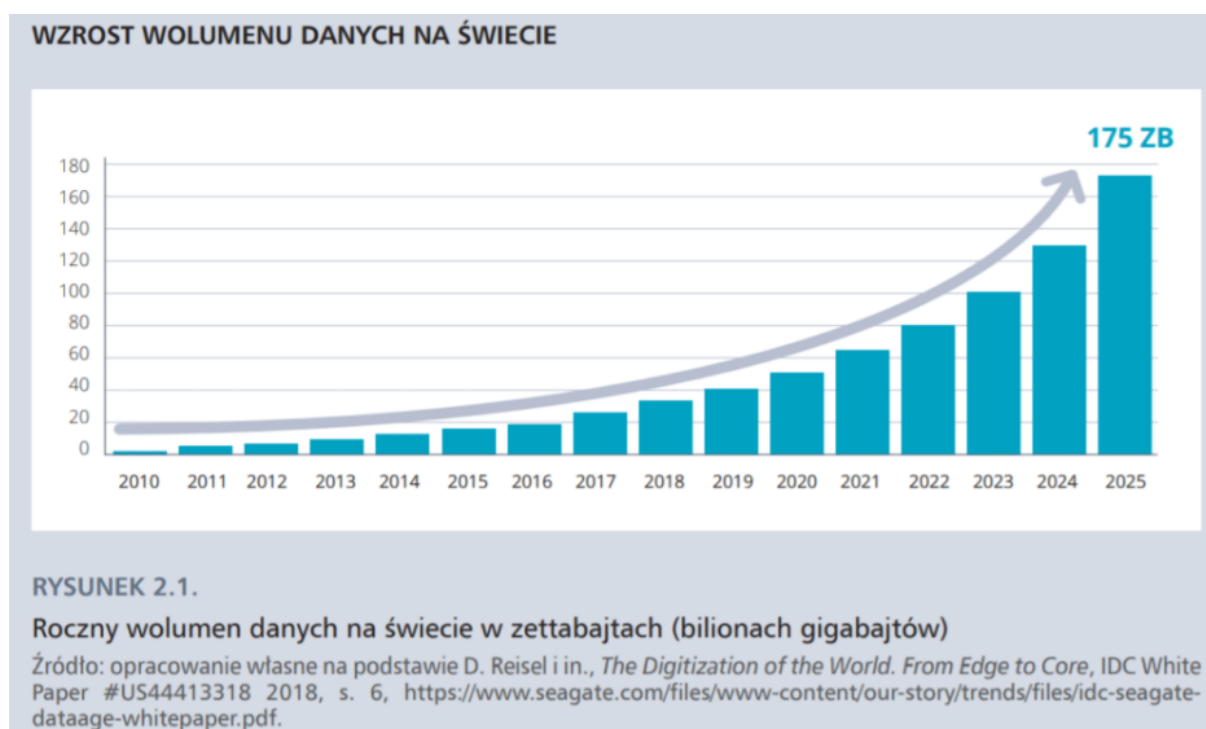


Data Science – uzupełnienie, czy ewolucja Controllingu?

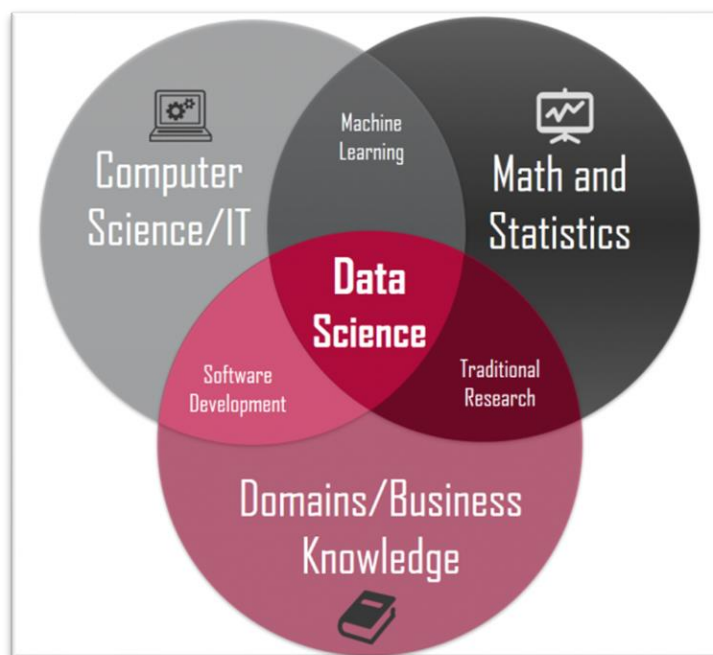
Niniejszy artykuł nawiązuje do webinaru zorganizowanego 22 kwietnia 2021 r. przez Grupę Roboczą ICV Szczecin. Podczas spotkania opowiedzieliśmy o tym, czym właściwie jest Data Science i dlaczego rola tej dziedziny rośnie, oraz dlaczego naszym zdaniem temat ten jest istotny dla dzisiejszego Controllera.

Dane gromadzone m.in. przez biznes rosną w tempie wykładniczym. Jest to związane z wieloma czynnikami, m.in. spadającymi cenami za przechowanie danych, rozwojem technologii w tym obszarze czy postępującą digitalizacją.



Te ogromne zbiory danych w organizacjach dają możliwość poszukiwania coraz nowszych driverów biznesu, czynników, które wpływają w sposób pośredni na wynik przedsiębiorstw czy kreowaną przez nie wartość dodaną. By przetwarzać taką liczbę danych konieczne jest używanie narzędzi wykraczających poza standardowy „toolbox” Controllera i tu z pomocą przychodzi szeroka dziedzina nazwana Data Science.

Według definicji Data Science to nauka o danych, to interdyscyplinarna dziedzina, która wykorzystuje naukowe metody, procesy, algorytmy i systemy do wydobywania wiedzy i spostrzeżeń z ustrukturyzowanych i nieustrukturyzowanych danych oraz do stosowania wiedzy i praktycznych wniosków z informacji w szerokim zakresie dziedzin aplikacji. Data Science jest związana z eksploracją danych, uczeniem maszynowym i big data (Wikipedia).



Powyższa definicja może się skojarzyć z zadaniami wykonywanymi przez statystyków czy speców od ryzyka, ludzi analizujących zjawiska z pomocą metod statystycznych i próbujących przewidywać ich kształtowanie w czasie. Jest to z resztą bardzo słuszne skojarzenie, ponieważ obecnie w tych zawodach również wykorzystywane są algorytmy uczenia maszynowego oparte o języki programowania, takie jak R czy Python oraz dedykowane oprogramowanie.

Controlling natomiast często kojarzy się z analizą historycznych wyników, próbą odpowiedzi na pytania – „Co było powodem osiągnięcia określonego wyniku/Jakie były najważniejsze przyczyny odchyień?”. W celu odpowiedzi na te pytania zaprężane są do pracy narzędzia analityczne, często również wykorzystujące pewne metody statystyczne.

Już na tym etapie uwydatnia się pewne podobieństwo – w obu przypadkach poszukiwane są zależności, główne czynniki wpływające na rezultat. Analityk ryzyka będzie poszukiwał odpowiedzi na pytanie: „jaka jest szansa niespłacenia kredytu przez określony profil klienta?”, Controller będzie odpowiadał „jakie zjawisko doprowadziło do ponadprzeciętnego odchylenia od budżetu?”. W tym drugim przypadku ważny jest element partnerstwa biznesowego, ale można go również wesprzeć wykorzystaniem metod statystycznych i uczenia maszynowego.

Czym jest zatem Machine Learning – jedno z głównych narzędzi Data Scientista oraz pojęcie, które w ostatnim czasie mocno zyskało na popularności, np. w wyszukiwarce Google:




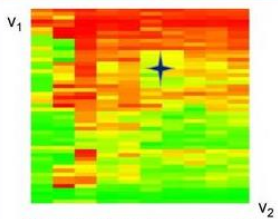
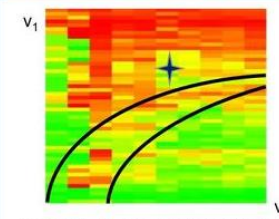
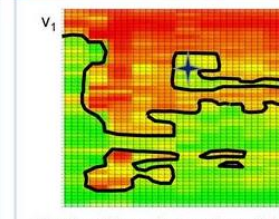
Źródło: Google trends

„Uczenie maszynowe (Machine Learning, ML) to nauka o algorytmach komputerowych, które poprawiają się automatycznie dzięki doświadczeniu i wykorzystaniu danych. Jest postrzegany jako część sztucznej inteligencji. Algorytmy uczenia maszynowego budują model na podstawie danych przykładowych, znanych jako „dane szkoleniowe”, w celu przewidywania lub podejmowania decyzji bez wyraźnego zaprogramowania do tego.” – Wikipedia

Różnicę między modelem ML a klasyczną statystyką trafnie ilustruje poniższy obraz przygotowany przez firmę McKinsey:

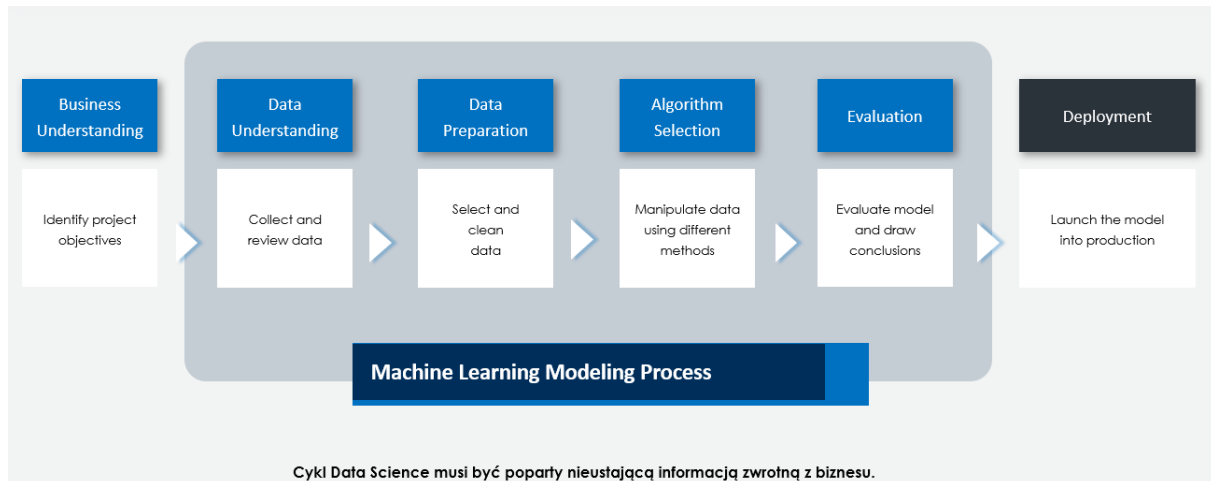
Machine learning surfaces insights within large, complex datasets, enabling more accurate risk models



The actual phenomenon (real historical data)	How Traditional stats see it	How Machine Learning sees it
 <p><i>Real life phenomenon come display complex, non linear patterns</i></p>	 <p><i>In simple regression techniques, the expert will pre-determine the range of transformations he/she sees as making business sense (linear, quadratic, etc.)</i></p>	 <p><i>ML algorithms, instead are able to tease out the right pattern, for the expert to sense check. Beware over-fitting and hypothesis-confirmation though !</i></p>
<p>Example: Decision to provide a mortgage to a specific client:</p> <ul style="list-style-type: none"> 25 years old, with masters degree Variable remuneration Fashion industry employee Rich family 	<p>Example: the client is wrongly identified as high-risk client through traditional stats</p> <p>✗ No mortgage provided</p>	<p>Example: the client is correctly identified as low-risk client through Machine Learning</p> <p>✓ Mortgage provided</p>

These algorithms “learn” with every bit of additional information as they identify new hidden patterns

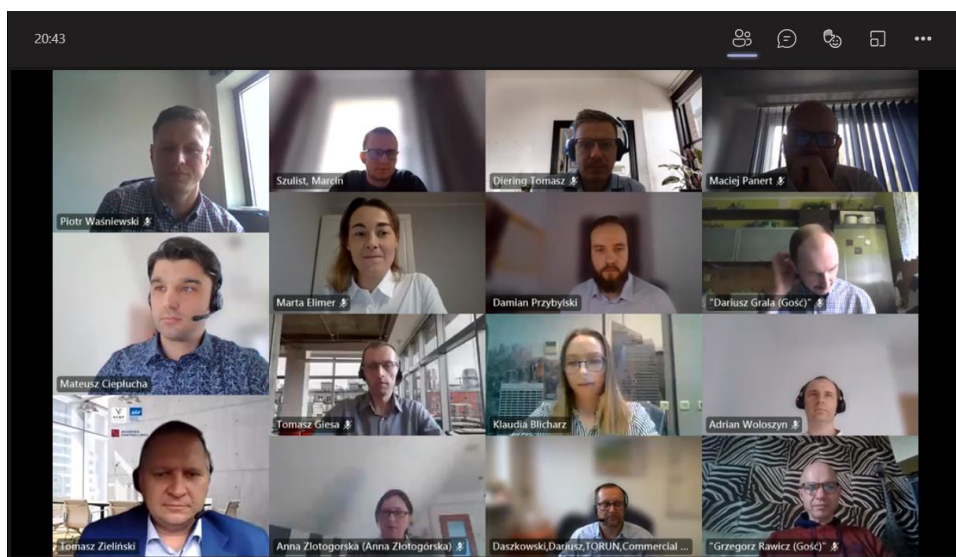
Wiemy zatem, że uczenie maszynowe polega docelowo na poprawie trafności prognoz pewnego zjawiska wraz ze stopniowym zdobywaniem przez model „doświadczenia” w postaci nowych danych. Przyjrzyjmy się więc jak powinien wyglądać modelowy proces wykorzystujący uczenie maszynowe w biznesie.

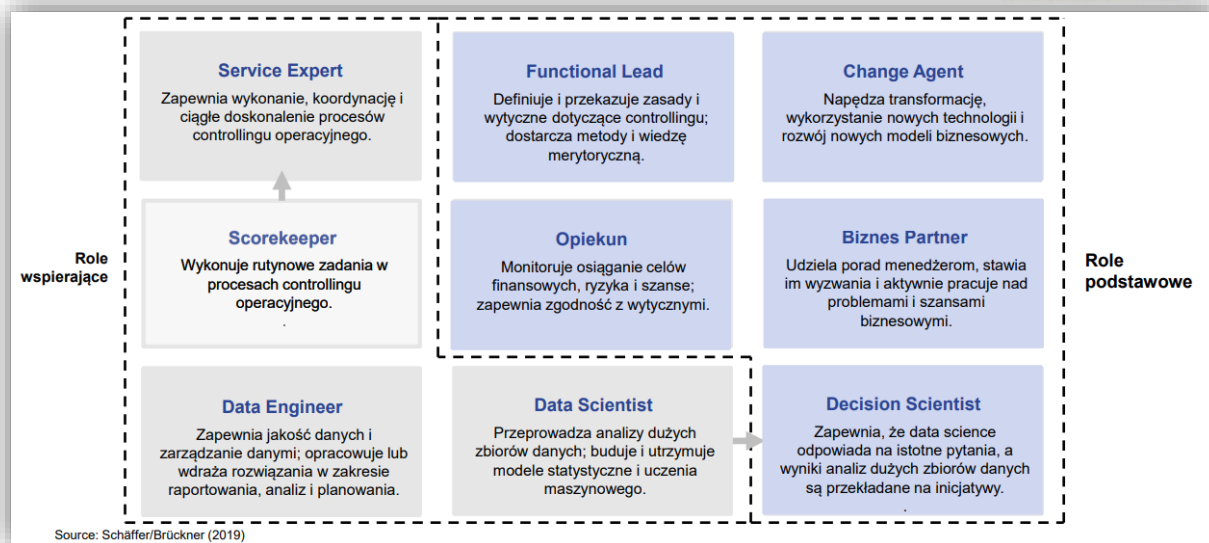


Źródło: opracowanie własne

Cały proces rozpoczyna się od zrozumienia problemu biznesowego, w czym przydatne może okazać się dobranie odpowiednich partnerów biznesowych. W kolejnych etapach przechodzi się do zrozumienia i obróbki danych w taki sposób, by mogły one zostać „skonsumowane” przez model. Zaprojektowane narzędzie podlega ewaluacji. Te etapy można odnieść do wzorcowej pracy Controlлера. Różnicą w tym przypadku jest rezultat – informacja zarządcza czy raport operacyjny w przypadku Controlлера, model predykcyjny w przypadku Data Scientista.

W porządku, mamy podobieństwa oraz różnice, ale co z tego wszystkiego wynika? Badania przeprowadzane przez Prof. Dr. Utza Schaeffera (Członek Zarządu ICV GLOBAL) wskazują, że umiejętności związane z Data Science są coraz bardziej potrzebne również w działach Controllingu.





W ramach szeroko pojętej pracy z danymi kształtują się nowe role. Obok partnera biznesowego będzie to również Decision Scientist, którego zadaniem jest w uproszczeniu dopasowanie narzędzi Data Science do potrzeb biznesu.

Są to możliwe kierunki rozwoju naszego zawodu, dlatego warto już teraz dodać omawiane w artykule narzędzia do swojego arsenału.

Opracował: Damian Przybylski – Lider Grupy Roboczej ICV Szczecin, Financial Controller w IAI S.A.