

3. EKSPLORACJA DANYCH (DATA MINING) I ODKRYWANIE WIEDZY (KNOWLEDGE DISCOVERY)

Autor: Dejan Mirčetić

Jak omówiono w Rozdziale 2, koncepcja ekstrakcji danych i generowania z nich wiedzy jest ściśle związana z **technikami Eksploracji Danych (ang. Data Mining techniques)**. We współczesnych przedsiębiorstwach techniki eksploracji danych wywierają duży wpływ na ogólną wydajność firm, ponieważ decyzje operacyjne, taktyczne i strategiczne są podejmowane w oparciu o informacje wejściowe uzyskane w procesie eksploracji danych. W Rozdziale 1 zauważono, że świat zawiera ponad 97 zettabajtów danych, a pojedyncze bazy danych często osiągają rozmiary terabajtów. W większości organizacji ilość danych podwaja się co dwa lata. Dane te są przechowywane na różnych platformach i w różnych formatach, zawierając dane ustrukturyzowane, nieustrukturyzowane i półustrukturyzowane (patrz Rozdział 1). Szacuje się, że do 90% danych biznesowych istnieje w formie nieustrukturyzowanej. Ponadto znaczna część tych danych może zawierać błędy wynikające z niewłaściwego przechowywania lub formatowania czy też błędów ręcznych powstających podczas zbierania danych. W rezultacie nie wszystkie dane przechowywane przez organizacje cechują się wymaganą dokładnością lub wiarygodnością. Niemniej jednak ta ogromna ilość danych zawiera cenne informacje strategiczne dla firm. Jednak w obliczu tak dużej ilości złożonych typów danych pojawia się pytanie W jaki sposób można je skutecznie „wydobyć”, aby uzyskać z nich istotne spostrzeżenia? Odpowiedź leży w eksploracji danych, która wspomaga zwiększanie przychodów i obniżanie kosztów poprzez szybkie i automatyczne wydobywanie przydatnej wiedzy i spostrzeżeń biznesowych z ogromnych zestawów danych.

Data Mining wyłoniło się z konieczności efektywnego wydobywania cennych informacji, wymagającego technik koncentrujących się na identyfikowaniu zrozumiałych wzorców, które można interpretować jako użyteczną lub interesującą wiedzę. Zatem **Data Mining jest iteracyjnym i interaktywnym procesem** mającym na celu odkrywanie ważnej, nowej, użytecznej i zrozumiałej wiedzy (wzorców, modeli, reguł itp.) w ogromnej bazie danych (Behera i in., 2019). Głównym celem Data Mining jest **ujawnienie krytycznych spostrzeżeń**, które **wspierają podejmowanie decyzji** w organizacji biznesowej.

W kolejnych podrozdziałach zajmiemy się tym, w jaki sposób dane są „wydobywane” na potrzeby analityki biznesowej i Odkrywania Wiedzy (ang. Knowledge Discovery).

3.1. Czym jest Eksploracja Danych (Data Mining)?

Przed zdefiniowaniem Data Mining ważne jest, aby umieścić to pojęcie w kontekście terminów, z którymi jest powszechnie kojarzone i powiązane oraz z tymi, z którymi jest często błędnie utożsamiany. Osoby niebędące ekspertami często mylą terminy Data Mining i Big Data technology. Są to jednak dwa odrębne pojęcia. **Big Data opisuje** niezwykle duże i złożone **zestawy danych**, które wymagają specjalistycznych aplikacji i oprogramowania do przetwarzania. Z drugiej strony **Data Mining idzie o krok dalej**, obejmuje analizę tak ogromnych ilości danych w celu odkrycia ukrytych reguł i wzorców, które mogą nie być widoczne na pierwszy rzut oka.

Data Mining to szerokie pojęcie obejmujące różne techniki analityczne, w tym statystykę, sztuczną inteligencję i uczenie maszynowe. Metody te służą do przeszukiwania ogromnych ilości danych przechowywanych w bazach danych organizacji lub repozytoriach online. Głównym celem eksploracji danych jest odkrywanie wzorców w zestawie danych. **Analityka Biznesowa (ang. Business Analytics – BA)** odnosi się do kompleksowego procesu wykorzystywania umiejętności, technologii, ustalonych praktyk i algorytmów związanych z Data Mining. **Stąd Data Mining powszechnie służy jako zaplecze funkcji BA**, podczas gdy front-end funkcji BA składa się z metryk raportowania dla kadry kierowniczej i zestawionych informacji przedstawionych w formacie, który umożliwia menedżerom podejmowanie świadomych decyzji biznesowych. Podczas korzystania z Data Mining profesjonaliści BA działają jak „detektywi danych” (Lee, 2013), analizując dane celem lepszego opisanie i zrozumienia obecnej i przeszłej sytuacji organizacji (analityka opisowa), przewidzenia przyszłych wyników (analityka predykcyjna) i podjęcia skutecznych działań (analityka preskryptywna).

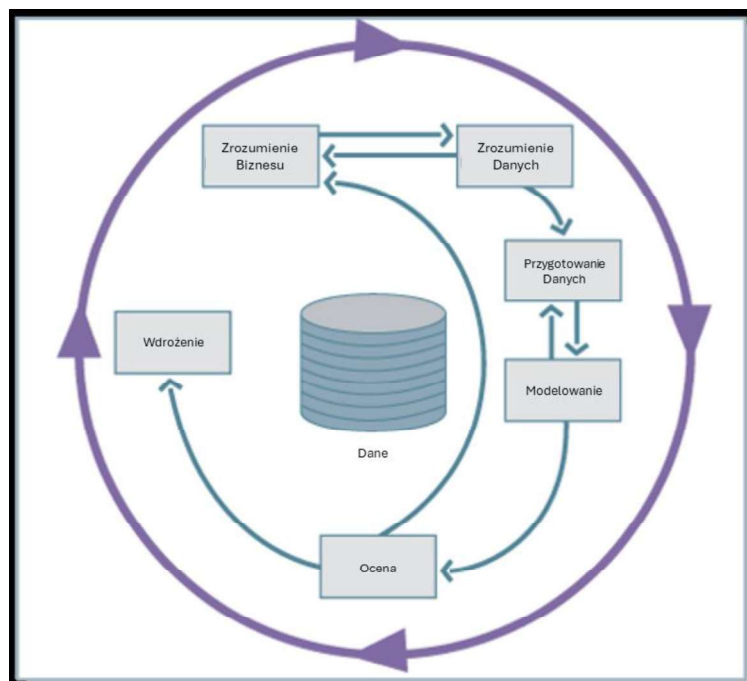
Data Mining jest podstawowym elementem procesu **Odkrywania Wiedzy w Bazach Danych (ang. Knowledge Discovery in Databases – KDD)**, ale stanowi tylko jeden krok w całym procesie. Aspekt Data Mining procesu KDD koncentruje się na wykorzystaniu algorytmów do ekstrakcji i identyfikacji wzorców z danych. W szerszym procesie KDD te wydobyte wzorce są oceniane i potencjalnie interpretowane w celu ustalenia, które wzorce można uznać za nową „wiedzę” (Behera i in., 2019). Zdefiniowane w ten sposób, z wykorzystaniem Data Mining jako zapleczem i Knowledge Discovery jako swego rodzaju

front-endem BA, wzorce te reprezentują, wraz z danymi biznesowymi, kluczowe filary analityki biznesowej, jak opisano w Rozdziale 2.

Data Mining wykorzystuje różnorodne algorytmy do eksploracji ogromnych zbiorów danych, identyfikując wzorce, które mogą dostarczać cenne informacje biznesowe. Data Mining to narzędzie, a nie magiczne rozwiązanie. Nie obserwuje biernie bazy danych organizacji i nie ostrzega jej zarządców o interesujących wzorcach. Zrozumienie danego przedsiębiorstwa, jego danych i metod analitycznych pozostaje kluczowe. Data Mining pomaga analitykom biznesowym odkrywać wzorce i relacje w danych, ale nie określa wartości tych wzorców dla organizacji. Dlatego Data Mining nie zastępuje wykwalifikowanych analityków biznesowych, zamiast tego zapewnia im potężne nowe narzędzie do usprawnienia ich pracy.

Data Mining obejmuje proces obliczeniowy polegający na identyfikowaniu trendów, reguł, ukrytych wzorców i innych cennych informacji poprzez analizę dużych zestawów danych. Data Mining dostosowuje swoją technikę korzystając z wielu obszarów badawczych, w tym statystyki, uczenia maszynowego, systemów baz danych, wizualizacji, sieci neuronowych itp. Jest to proces wydobywania praktycznej wiedzy z różnych źródeł danych przechowywanych w zróżnicowanych formatach. Data Mining stał się coraz bardziej istotny w ostatnich latach ze względu na postęp obserwowany w technologiach przechowywania danych (ang. Big Data), sztucznej inteligencji (ang. Artificial Intelligence – AI) i automatyzacji procesów robotycznych (ang. Robotic Process Automation – RPA).

Proces Odkrywania Wiedzy w Bazach Danych (KDD) obejmuje wykorzystanie bazy danych, w tym dokonania niezbędnego wyboru, wstępnego przetwarzania, wstępnego próbkowania i transformacji, w celu zastosowania algorytmów Data Mining do identyfikacji wzorców (Behera i in., 2019). Obejmuje on również ocenę wyników Data Mining. Wspólnym standardem opisu kroków procesu Knowledge Discovery jest Międzybranżowy Standardowy Proces Eksploracji Danych (ang. Cross-Industry Standard Process for Data Mining – CRISP-DM), który pokazano na Rysunku 3.1.



Rysunek 3.1 Międzybranżowy Standardowy Proces Eksploracji Danych (ang. Cross-Industry Standard Process for Data Mining – CRISP-DM)

Źródło: Rahman i in. (2016)

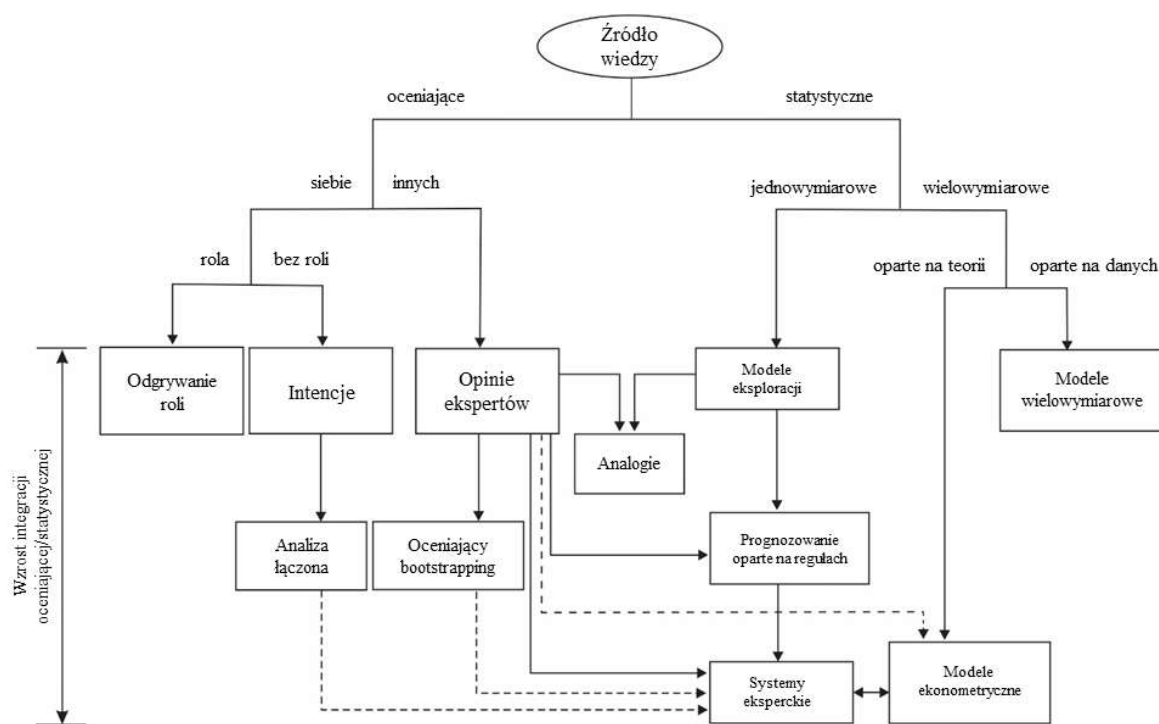
Na Rysunku 3.1 pierwszą fazę procesu stanowi zrozumienie biznesowe, obejmujące zrozumienie celów, które mają zostać osiągnięte, a także przeprowadzenie szczegółowego ustalania faktów na temat zasobów i założeń. Druga faza, zrozumienie danych, skupia się na badaniu różnych opisowych charakterystyk danych. Trzecia faza, przygotowanie danych, jest najbardziej wymagającą i czasochłonną częścią procesu KDD, mającą na celu wybranie odpowiednich danych i odpowiednie sformatowanie ich do analizy. Ta faza obejmuje działania takie jak wybór danych, filtrowanie, transformacja i integracja. Czwarta faza, modelowanie, obejmuje stosowanie metod analitycznych i wybór najbardziej odpowiednich algorytmów. Ta faza obejmuje również weryfikację jakości modelu poprzez testowanie i walidację krzyżową (sprawdzian krzyżowy). Piąta faza, ocena, obejmuje interpretację i ocenę odkrytej wiedzy (Rahman i in., 2016).

3.2. Odkrywanie Wiedzy w Logistyce i Zarządzaniu Łańcuchem Dostaw

W kontekście ŁD i logistyki Data Mining różni się od innych aplikacji ogólnego przeznaczenia. Powód jest związany z wspomnianą i omówioną wcześniej różnorodnością źródeł i struktury danych biznesowych, które stanowią duże wyzwanie dla inżynierów

projektujących odpowiednie rozwiązania modelowe. Proces generowania wiedzy z danych można podsumować na Rysunku 3.2.

Na podstawie Rysunku 3.2, odkrywanie i generowanie wiedzy z danych jest w dużym stopniu zależne od źródła wiedzy i można je podzielić na **oceniające** i **statystyczne**. Oba te kierunki mają swoje zalety, ale procedura ekstrakcji wiedzy ze źródła jest zasadniczo różna. Aby zastosować bardziej formalne podejście do eksploracji danych, tj. podejście statystyczne, kluczowym fundamentem jest istnienie danych ilościowych. W łańcuchach dostaw istnieje duża liczba sektorów, lokalizacji i transakcji, generujących przepływy danych, które można wykorzystać do eksploracji i ekstrakcji przydatnych informacji zwrotnych. Z drugiej strony, w ŁD i logistyce istnieje również wiele źródeł danych, które nie mają charakteru ilościowego, a zatem nie podlegają formalnym procedurom ilościowym. Zamiast tego dane te są tradycyjnie poddawane eksperckim panelom oceniającym (metoda delficka), a decyzje są podejmowane na podstawie doświadczenia, wiedzy i autorytetu ekspertów.



Rysunek 3.2 Zasady ekstrakcji wiedzy z danych

Źródło: Armstrong (Ed.) (2001)

W tej książce nacisk zostanie położony na **techniki ilościowe/matematyczne**, choć krótko omówimy również niektóre metodologie przechwytywania wiedzy eksperckiej w ustrukturyzowanych ramach, tj. omówimy Systemy Ekspertskie (ang. Expert Systemes – ES) i ich zastosowania oraz przykłady.

3.3. Podejście Delfickie do Oceniającego Tworzenia Wiedzy

Cenne spostrzeżenia doświadczonych profesjonalistów w dziedzinie łańcucha dostaw i logistyki często pozostają niezauważone. Te spostrzeżenia należy udostępniać mniej doświadczonym profesjonalistom w tej dziedzinie. Metoda delficka jest jednym ze sposobów pozyskiwania i rozpowszechniania wiedzy eksperckiej. Według (Steurer, 2011) metoda delficka, nazwana na cześć wyroczni delfickiej, która początkowo była wykorzystywana do konsultacji w różnych sprawach publicznych i osobistych w starożytnej Grecji, w latach 50. XX wieku przekształciła się w technikę, w której eksperci zastępowali wyrocznię, aby osiągnąć konsensus wśród grupy ekspertów w danej dziedzinie. „Project Delphi”, finansowany przez Siły Powietrzne USA, był pierwszym projektem wykorzystującym tę metodę do prognozowania rozwoju technologicznego. Od tego czasu metoda delficka ewoluowała i udoskonalała się, znajdując zastosowania w różnych dyscyplinach naukowych. Metoda delficka została opracowana w celu osiągnięcia wiarygodnego konsensusu ekspertów, często służąc jako substytut dowodów empirycznych, gdy takich dowodów brakuje. Oznacza to, że technika delficka jest procesem iteracyjnym, w którym eksperci anonimowo wydają osądy na konkretny temat, mając na celu osiągnięcie konsensusu i sprzeciwu wraz z ich uzasadnieniami. Jest to wysoce ustrukturyzowany proces komunikacji grupowej, w którym eksperci oceniają niepewną i niekompletną wiedzę (Naisola-Ruiter, 2022). Według Paivarinty i in. (2011) metoda delficka, między innymi, jest szeroko wykorzystywana w badaniach nad systemami informacyjnymi. Jest stosowana do wybierania projektów w zakresie Systemów Informacyjnych (ang. Information Systems – IS), ustalania priorytetów ryzyka projektów rozwoju oprogramowania, definiowania wymagań projektu IS, określania kluczowych kwestii w zarządzaniu IS, tworzenia ram dla działań związanych z manipulacją wiedzą, zrozumienia ról i zakresu systemów zarządzania wiedzą w organizacjach oraz monitorowania badań nad systemami informacyjnymi w zakresie offshoringu (przeniesienia wybranych procesów działalności poza granicę kraju).

Metoda delficka stała się standardową praktyką w zakresie ilościowego określania wyników procesów pozyskiwania uwagi w grupie. Jest wykorzystywana w różnych dyscyplinach do prognozowania trendów, ustalania priorytetów obszarów badawczych, oceny potencjalnych skutków różnych wyborów politycznych, ustalania wskaźników wydajności i opracowywania wytycznych klinicznych. **Techniki delfickie są również stosowane w obszarze ŁD i logistyki.** Na przykład jest ona wysoce zalecana jako instrument do identyfikacji i oceny ryzyka dostaw, do różnego rodzaju oceny w procesach logistycznych, do określania najlepszych praktyk logistycznych, do strategicznego podejmowania decyzji i opracowywania polityki

przedsiębiorstwa, do mapowania przyszłych praktyk SCM i do prognozowania logistyki. Cztery kluczowe cechy lub podstawowe zasady metody delfickiej to:

- Proces iteracyjny i wieloetapowy (oraz zbieranie danych);
- Opinie uczestników (kontrolowane na pewnym poziomie) z możliwością poprawienia swoich odpowiedzi
- Statystyczne określenie odpowiedzi grupy;
- Pewien stopień anonimowości.

Typowy proces delficki obejmuje przedstawienie serii pytań w wielu rundach. Paneliści, wybrani ze względu na swoją wiedzę i doświadczenie, odpowiadają anonimowo. Po każdej rundzie następuje informacja zwrotna na temat zagregowanych odpowiedzi, pozwalając uczestnikom dostrzec, jak kształtują się ich odpowiedzi w kontekście odpowiedzi całego panelu. Paneliści mogą następnie dostosować swoje odpowiedzi i podać uzasadnienie wszelkich zmian w kolejnych rundach. Ten iteracyjny proces trwa do momentu osiągnięcia konsensusu lub ukończenia ustalonej liczby rund.

3.3.1. Kroki przeprowadzania metody delfickiej

Metoda delficka to ustrukturyzowane podejście, które polega na zbieraniu opinii i spostrzeżeń ekspertów w celu osiągnięcia konsensusu w danym temacie. Proces ten zazwyczaj obejmuje cztery główne kroki (Rysunek 3.3).



Rysunek 3.3 Kroki przeprowadzania metody delfickiej

Źródło: Steurer (2011)

Krok 1 – Określenie celów: Pierwszy krok obejmuje określenie celów i zakresu badania delfickiego. Krok ten zawiera więc identyfikację konkretnych pytań lub tematów wymagających wkładu eksperta i nakreślenie kluczowych kwestii do omówienia. Ten podstawowy krok zapewnia, że badanie zachowuje koncentrację i celowość przez cały czas jego trwania.

Krok 2 – Wybór ekspertów: Wybór właściwej grupy ekspertów jest niezbędny dla sukcesu i skuteczności techniki delfickiej. Eksperci ci powinni posiadać odpowiednią wiedzę, umiejętności i doświadczenie dotyczące badanego tematu. Grupa powinna być zróżnicowana,

aby zapewnić szerokie spektrum punktów widzenia. Liczba uczestników może się różnić w zależności od wielkości i złożoności badania, ale ogólnie zaleca się uwzględnienie co najmniej 10–15 ekspertów.

Krok 3 – Opracowanie i uruchomienie kwestionariuszy: Ten krok obejmuje opracowanie kwestionariuszy w celu zebrania opinii od ekspertów. Zazwyczaj początkowy kwestionariusz jest otwarty, aby umożliwić ekspertom swobodne dzielenie się swoimi opiniami bez wpływu z zewnątrz. W rundzie 1 eksperci otrzymują kwestionariusz otwarty i niezależnie przekazują spostrzeżenia, przewidywania lub sugestie związane z celami badania. W rundzie 2 moderator podsumowuje i anonimizuje odpowiedzi z rundy 1, aby utworzyć bardziej ukierunkowany kwestionariusz na następną rundę. Jeśli zajdzie taka potrzeba, można przeprowadzić dodatkowe rundy w celu dopracowania opinii na podstawie osiągniętych poziomów konsensusu, kontynuując do momentu osiągnięcia wstępnie zdefiniowanego konsensusu lub podjęcia przez moderatora decyzji o zakończeniu procesu.

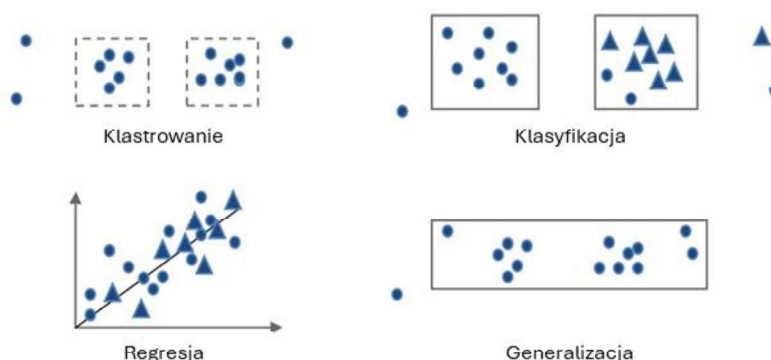
Krok 4 – Wykorzystanie wyników: Po zakończeniu procesu delfickiego i osiągnięciu konsensusu, wyniki są analizowane i wykorzystywane do podejmowania decyzji, prognozowania, opracowywania polityki organizacji lub innych celów określonych w kroku 1. Anonimowość badań delfickich pomaga zapewnić, że końcowe wyniki są bezstronne i odzwierciedlają łączną wiedzę ekspertów zaangażowanych w badanie.

3.4. Podejście Ilościowej Eksploracji Danych do Odkrywania Wiedzy

Jak pokazano na Rysunku 3.2, ilościowa **Eksploracja Danych jest mocno wspierana przez formalne narzędzia matematyczne, a bardziej bezpośrednio statystyczne**. Możemy tutaj użyć argumentu, że każda operacja statystyczna oparta na danych może być uważana za analizę ilościową. Głównym celem tych operacji jest wyodrębnienie rzeczywistych wzorców z danych i wygenerowanie przydatnych spostrzeżeń w obserwowanym procesie (w przypadku analizy w biznesie lub łańcuchach dostaw). Nie jest to proste zadanie, ponieważ istnieje znacząca niezgodność między założeniami teorii statystycznej/matematycznej a rozkładami i wzorcami, które są obecne w rzeczywistych danych biznesowych. W praktyce główny przepływ większości analiz biznesowych oparty jest właśnie na tej niezgodności. Stosując metody ilościowe niezwykle ważne jest, aby dane spełniały teoretyczne założenia matematyczne ograniczone przez obserwowany model, tak aby

można było traktować wyniki modelu jako ważne i potencjalnie podejmować decyzje na ich podstawie.

Jak omówiono w poprzedniej sekcji, metody eksploracji danych stanowią podstawowy komponent procesu KDD i są wielokrotnie w nim wykorzystywane. Eksploracja danych jest techniką multidyscyplinarną, której ostateczne metody analityczne opierają się na matematyce. Statystyka odgrywa szczególnie istotną rolę w procesie analizy danych w fazie przygotowywania danych, stanowiąc podstawę dla kilku metod eksploracji danych. Eksploracja danych obejmuje wykorzystanie wydajnych algorytmów w celu odkrycia oczekiwanych lub domniemanych wzorców. Jak pokazano na Rysunku 3.4, zadania Eksploracji Danych można podzielić na pięć kategorii: klastrowanie, klasyfikacja, regresja, reguły asocjacyjne i generalizacja. **Klastrowanie** ma na celu grupowanie obiektów bazy danych tak, aby obiekty w obrębie klastra były do siebie podobne, podczas gdy obiekty w różnych klastrach różniły się od siebie. **Klasyfikacja** obejmuje naukę funkcji, która przypisuje wartości atrybutów do wstępnie zdefiniowanych klas. **Regresja**, metoda statystyczna szacująca relacje między zmiennymi, powszechnie stosowana do przewidywania i prognozowania, znacznie pokrywając się z uczeniem maszynowym. **Reguły asocjacyjne** są stosowane do opisywania silnych relacji w procesach transakcyjnych, takich jak „gdy A i B, to C”. **Generalizacja** ma na celu wyrażenie dużej ilości danych w możliwie jak najbardziej zwarty sposób (Su, 2016). Główne techniki stosowane w Data Mining to: reguły klasyfikacji lub drzewa decyzyjne, regresja, klastrowanie, algorytmy genetyczne, modelowanie oparte na agentach itp.

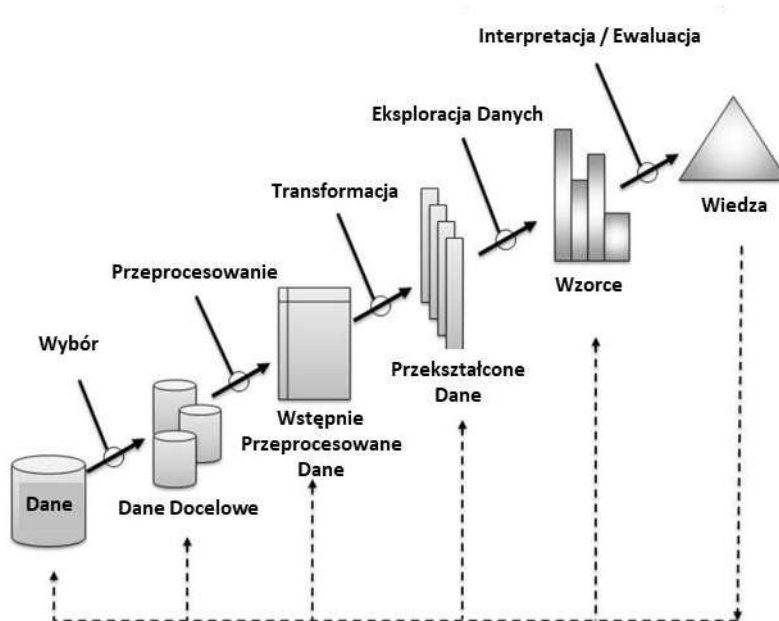


Rysunek 3.4 Zadania Eksploracji Danych

Source: Su (2016)

Wiedzę można wyodrębnić z przetworzonych danych ilościowych. **Proces KDD obejmuje dziewięć kroków** przedstawionych na Rysunku 3.5. Ważne jest, aby zauważyć, że eksploracja danych jest przeprowadzana na bazie przekształconych danych, z których uprzednio wykluczono nieistotne informacje. Wzory odkryte w tym procesie są następnie

interpretowane i oceniane w określonym kontekście w celu uzyskania wiedzy, która może wspomóc proces decyzyjny.



Rysunek 3.5 Kroki składające się na proces KDD

Źródło: Fayyad i in. (1996)

Jak już wcześniej wspomniano, powszechnym standardem opisu etapów procesu KDD jest CRISP-DM, który stanowi wiodący model przemysłowy. W odniesieniu do bieżących publikacji i badań przeanalizowanych przez Su (2016), typowe techniki eksploracji danych i zastosowania w łańcuchach dostaw obejmują:

- **Drzewa decyzyjne** (Zastosowanie: rozwiązywanie problemów dostawców, które można sprowadzić np. do zestawu możliwych wyników dla każdej decyzji, wraz z oceną prawdopodobieństwa wystąpienia każdego wyniku).
- **Regresję** (Zastosowanie: prognozowanie i szacowanie popytu klientów na nowy produkt).
- **Regułę asocjacyjną** (Zastosowanie: identyfikowanie przyczyn awarii produktu, optymalizacja zdolności produkcyjnej i umożliwienie utrzymania ruchu opartego na stanie jakościowym).
- **Algorytm genetyczny** (Zastosowanie: rozwijanie hipotezy poprawy działania VMI (Zarządzania Zapasami Przez Dostawcę – ang. Vendor Managed Inventory) w niepewnym środowisku popytu).

- **Algorytmy klastrowania** (Zastosowanie: algorytm centroidów (k-średnich – ang. k-Mean) do kategoryzacji zwracanych towarów w celu poprawy jakości procesów produkcyjnych).
- **System eksploracji danych wieloagentowych** (Zastosowanie: wspieranie decyzji dotyczących planowania produkcji w oparciu o analizę historycznych danych o popycie na produkty).

Wiedza wydobyta z wykorzystaniem Data Mining jest zazwyczaj przechowywana i prezentowana przy użyciu **Systemów Eksperskich (ang. Experts Systems – ES)**. ES to wyrafinowany system wiedzy zaprojektowany w celu naśladowania ludzkiej wiedzy specjalistycznej w różnych obszarach zastosowań. Olson i Courtney (1992) definiują ES jako program komputerowy w określonej domenie, obejmujący pewną ilość sztucznej inteligencji w celu naśladowania ludzkiego myślenia w drodze dojścia do tych samych wniosków, co ekspert danej dziedziny. Komponent ES jest idealny do wspomagania decydenta w obszarze, w którym wymagana jest wiedza specjalistyczna (Turban, 1995). Zasadniczo ES przenosi wiedzę specjalistyczną od eksperta (lub innego źródła) do komputera. Może on albo wspierać decydentów, albo całkowicie ich zastępować i jest najszerzej stosowaną, a także najbardziej udaną komercyjnie technologią sztucznej inteligencji (Turban i in. 2007). Jednym z uzasadnień budowy ES jest dostarczanie wiedzy eksperckiej dużej liczbie użytkowników (Kock, 2005). Według Turbana i in. (2007), ES są uważane za część Systemów Wsparcie Decyzji (ang. Decision Support System – DSS), który można scharakteryzować jako komputerowy system informacyjny łączący modele i dane w celu rozwiązywania problemów półstrukturalnych i niestrukturalnych z wysokim poziomem zaangażowania użytkownika (Mirčetić i in. 2016, Turban i in. 2007). Następny rozdział omawia uczenie maszynowe (ang. Machine Learning – ML), które odnosi się do zdolności komputerów do uczenia się i reprezentowania wiedzy pochodzącej z danych wejściowych do procesu. ML można postrzegać jako pomost między wynikami uzyskanymi przez Data Mining a narzędziami Business Intelligence używanymi do prezentowania metryk raportowania dla kadry kierowniczej w formacie umożliwiającym menedżerom podejmowanie świadomych decyzji biznesowych.

BIBLIOGRAFIJA

1. Behera, P. C., Dash, C. & Mohapatra, S. (2019). Data Mining and Knowledge Discovery (KDD). *International Journal of Research and Analytical Reviews*, 6(1), s. 101-106.
2. Fayyad, U. M., Patetsky-Shapiro, G. & Smith, P. (1996). From Data Mining to Knowledge Discovery in Databases. *American Association for Artificial Intelligence*, 17, s. 37-34.
3. Kock, E. D. (2005) Decentralising the Codification of Rules in a Decision Support Expert Knowledge Base (MSc thesis). University of Pretoria; 2005. Dostępne: <http://repository.up.ac.za/handle/2263/22959>
4. Lee, P. M. (2013). Use of Data Mining in Business Analytics to Support Business Competitiveness. *Review of Business Information Systems*, 17(2), s. 53-58.
5. Mirčetić, D., Ralević, N., Nikoličić, S., Maslarić, M. & Stojanović, Đ. (2016). Expert system models for forecasting forklifts engagement in a warehouse loading operation: A case study. *Promet-Traffic&Transportation*, 28(4), s. 393-401.
6. Naisola-Ruiter, V. (2022). The Delphi technique: a tutorial. *Research in Hospitality Management*, 12(1), s. 91-97.
7. Olson, D. L., Courtney, J. F. & Courtney, J. F. (1992). *Decision support models and expert systems*. New York: Macmillan, USA.
8. Paivarinta, T., Pekkola, S. & Moe, C. E. (2011). Grounding Theory from Delphi Studies. In: *Proceedings of the 32nd International Conference on Information Systems (ICIS 2011): Research Methods and Philosophy*, 4-7 December 2011, Shanghai, China.
9. Rahman, F. A., Shamsuddin, S. M., Hassan, S. & Haris, N. A. (2016). A Review of KDD-Data Mining Framework and Its Application in Logistics and Transportation. *International Journal of Supply Chain Management*, 2(1), s. 1-9.
10. Steurer, J. (2011). The Delphi method: an efficient procedure to generate knowledge. *Skeletal Radiol*, 40, s. 959-961.
11. Su, W. (2016). Knowledge Discovery in Supply Chain Transaction Data by Applying Data Farming (Master thesis). Technical University of Dortmund, Dortmund, DE.
12. Turban, E. (1995). *Decision support and expert systems Management support systems*. Prentice-Hall, Inc. New York, USA.

13. Turban, E., Rainer, R. K. & Potter, R. E. (2007). Introduction to Information Systems: Supporting and Transforming Business. John Wiley & Sons, Inc. USA.