

Konrad KULIKOWSKI
Uniwersytet Jagielloński
Instytut Psychologii
konrad.kulikowski@uj.edu.pl

ZASTOSOWANIE MODELU CROSS INDUSTRY STANDARD PROCESS FOR DATA MINING (CRISP-DM) W BADANIACH POSTAW I OPINII PRACOWNIKÓW

Streszczenie. Celem artykułu jest przedstawienie modelu Cross Industry Standard Process for Data Mining (CRISP-DM) jako kompleksowego modelu zbierania i analizy danych w badaniach postaw i opinii pracowników. Model CRISP-DM przez ustrukturyzowanie i porządkowanie procesu badania może usprawnić zarządzanie nim, a także umożliwić efektywniejsze odkrywanie wiedzy ze zgromadzonych danych.

Słowa kluczowe: postawy, opinie, CRISP-DM, badania pracowników, analiza danych.

APPLYING CROSS INDUSTRY STANDARD PROCESS FOR DATA MINING (CRISP-DM) FOR EMPLOYEES ATTITUDES AND OPINION RESEARCH

Summary. The aim of this study is to present a Cross Industry Standard Process for Data Mining (CRISP-DM), as a model of collecting and analyzing data from employees attitudes and opinions research. CRISP – DM model through structuring and organization of the research process can improve research management and enable more efficient knowledge discovery from collected data.

Keywords: attitudes, opinions, CRISP-DM, employee surveys, data analysis.

1. Wprowadzenie

Postawy i przekonania pracowników to jedne z częściej analizowanych zagadnień z zakresu psychologii organizacji [11]. Badania wskazują, że pozytywne postawy pracowników [13, 22] oraz ich zaangażowanie [3, 9] są istotnymi czynnikami sukcesu przedsiębiorstwa. Firmy, które w systematyczny sposób pytają pracowników o zdanie, obserwują istotny wzrost poziomu wykonania pracy [19]. Postawy wobec pracy i organizacji wiążą się z rotacją, liczbą spóźnień [8], nieobecnościami w pracy [7], jakością świadczonych usług [15] oraz poziomem sprzedaży [26]. Nie bez znaczenia są też opinie pracowników o przełożonych [24], zwłaszcza przekonanie, że przełożeni są uczciwi [5,6]. Istotne jest też poczucie pracownika, że pasuje on do specyfiki organizacji, w której jest zatrudniony [18, 20].

Analiza postaw i opinii pracowników może dostarczyć cennej wiedzy o funkcjonowaniu i problemach organizacji, wciąż jednak wyzwaniem pozostaje zarządzanie procesem badania i skuteczne pozyskiwanie wiedzy z zebranych danych. Celem artykułu jest przedstawienie możliwości wykorzystania metodyki Cross Industry Standard Process for Data Mining (CRISP-DM) jako kompleksowego modelu prowadzenia badań postaw i opinii pracowników.

2. Cross Industry Standard Process for Data Mining, CRISP-DM

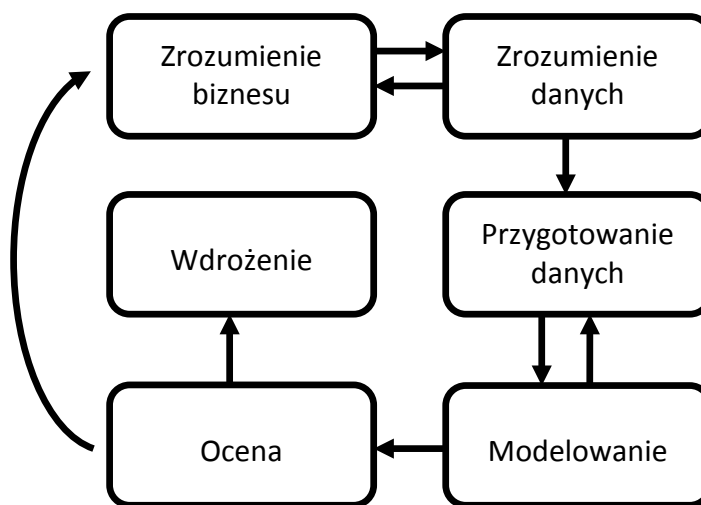
If you torture the data long enough, Nature will confess

Ronald Coase

Cross Industry Standard Process for Data Mining (CRISP-DM) to model opisujący proces odkrywania wiedzy z danych, opracowany przez organizacje biznesowe, zainteresowane odnośnieniem praktycznych i realnych korzyści z eksploracji danych [28]. Model ten systematyzuje wskazówki, działania i dobre praktyki, mające zastosowanie w eksploracji danych [4]. Model CRISP-DM zakłada [14], że odkrywanie wiedzy z danych to proces, którego cykl życia składa się z sześciu etapów: (1) zrozumienia uwarunkowań biznesowych, (2) zrozumienia danych, (3) przygotowania danych, (4) modelowania/analizy danych, (5) oceny rezultatów oraz (6) wdrożenia (por. rys. 1). W ramach każdego z etapów zdefiniowane są zadania, które należy zrealizować (por. tab. 1). Model ten zakłada, że zbieranie i analiza danych nie mogą być prowadzone bez zrozumienia kontekstu funkcjonowania organizacji. Wskazuje także, że ostatecznym celem analizy danych jest pozyskanie wiedzy pozwalającej na wdrożenie praktycznych rozwiązań [2]. Założeniem modelu jest także płynność przechodzenia pomiędzy etapami. Wiedza pozyskana na aktualnym etapie może zmieniać

wcześniejsze założenia i umożliwić ich przeformułowanie. Zawsze bierze się pod uwagę możliwość powrotu do etapu wcześniejszego.

Wytyczne CRISP-DM znajdują szerokie zastosowanie w obszarze analizy danych. Xiao i in. [29] wskazują przykładowo, że jest to odpowiedni sposób nauczania eksploracji danych, który można stosować w edukacji przyszłych analityków. Model ten znajduje zastosowanie w kryminalistyce [27], zarządzaniu finansami [10] czy w ocenie efektywności oddziaływań edukacyjnych [1]. Z sukcesami podejmuje się próby wdrażania założeń CRISP-DM w analizie wyników badań klinicznych [12, 17, 23]. Shearer [25] zauważa, że CRISP-DM jest popularny, ponieważ bazuje na solidnych, empirycznych doświadczeniach i był tworzony z myślą o praktycznych zastosowaniach.



Rys. 1. Kolejne etapy eksploracji danych według modelu Cross Industry Standard Process for Data Mining CRISP-DM

Fig. 1. The steps of data mining process according to Cross Industry Standard Process for Data Mining CRISP -DM

Źródło: opracowanie własne na podstawie: Chapman P., Clinton J., Kerber R., Khabaza T., Reinartz T., Shearer, C., (2000) CRISP-DM 1.0: Step-by-step data mining guide, CRISP-DM Consortium.

3. Model CRISP-DM jako model prowadzenia badań postaw i opinii pracowników

CRISP-DM ze względu na wskazywaną w literaturze użyteczność praktyczną [16, 21] może służyć jako kompleksowy zbiór wskazówek i metod postępowania w procesie analizy postaw i opinii pracowników. Stosowanie założeń CRISP-DM pozwala na planowanie, monitorowanie i kontrolowanie przebiegu badania opinii. Wyróżnienie etapów oraz określenie w ramach każdego z nich zadań do realizacji systematyzuje proces badawczy. CRISP-DM tworzy swoistą mapę drogową ukazującą, jakie działania należy podjąć w danym

momencie, przez co pozwala odejść od podejścia „zróbmy coś z tymi danymi” na rzecz systematycznej i uporządkowanej analizy. Dzięki właściwościom klaryfikującym i porządkującym, CRISP-DM zwiększa szansę uzyskania z danych wiedzy, pozwalającej na praktyczne działania [25]. Model systematyzujący przebieg badania postaw pracowniczych może być użyteczny zarówno dla analityków danych, jak i osób niemających doświadczenia w analizie, a uczestniczących w procesie badania opinii pracowników. Szczegółowy opis modelu przedstawił Chapman i in. [2], natomiast poniżej zwięźle opisano etapy i zadania CRISP-DM w kontekście badania postaw i opinii pracowników. Jest to nie tyle adaptacja – gruntowna zmiana modelu w celu przystosowania go do innych zastosowań, ile próba zastosowania istniejącego modelu w kontekście, w którym nie był dotychczas wykorzystywany. Takie podjęcie tematu uzasadniać może fakt, że model CRISP-DM jest z założenia modelem heurystycznym o szerokiej klasie zastosowań. Ponadto, istniejący model zyskuje potwierdzenie w badaniach empirycznych, zatem celowe wydaje się raczej szukanie nowych zastosowań dla istniejącego, sprawdzonego w praktycznym działaniu modelu, niż próby modyfikowania go, mogące wiązać się z utratą rzetelności. Dotychczasowe możliwości aplikacyjne modelu CRISP-DM poszerzone tu zostają przez określenie nowego obszaru rzeczywistości, w którym model można aplikować, co biorąc pod uwagę popularność badań postaw pracowniczych wydaje się być znacznym rozszerzeniem jego dotychczasowej użyteczności.

Tabela 1

Poszczególne etapy modelu CRISP-DM wraz z przypisanymi im zadaniami

Etap 1. Zrozumienie uwarunkowań biznesowych	Etap 2. Zrozumienie danych
Określenie celów biznesowych organizacji	Zebrańie wstępnych danych
Ocena kontekstu prowadzenia analizy w organizacji	Stworzenie opisu danych
Określenie celów analizy danych	Wstępna eksploracja danych
Stworzenie planu analizy danych	Weryfikacja jakości danych
Etap 3. Przygotowanie danych	Etap 4. Modelowanie
Wybór danych do analizy	Wybór techniki modelowania/analizy danych
Czyszczenie danych	Stworzenie kryteriów testowania wyników
Konstruowanie nowych zmiennych	Budowa model
Formatowanie i scalanie danych	Ocena modelu
Etap 5. Ewaluacja	Etap 6. Wdrożenie
Ewaluacja rezultatów	Plan wdrożenia
Przegląd procesu analizy	Plan monitorowania i utrzymania działań
Wdrożenie działań lub ponowna analiza	Raport końcowy
Raport z analizy	Przegląd projektu

Źródło: Opracowanie własne na podstawie: Chapman P., Clinton J., Kerber R., Khabaza T., Reinartz T., Shearer C. (2000) CRISP-DM 1.0: Step-by-step data mining guide, CRISP-DM Consortium.

Głównym warunkiem leżącym u podstaw korzystania z modelu CRISP-DM, w kontekście badań opinii pracowniczych, jest zebranie rzetelnych, trafnych i reprezentatywnych danych o istotnych dla pracowników aspektach funkcjonowania organizacji. Drugim kluczowym warunkiem wykorzystania modelu jest poprawne przeprowadzenie etapu analizy uwarunkowań biznesowych organizacji. Określenie na tym etapie celów biznesowych organizacji, kontekstu prowadzenia analizy, celów analizy danych oraz stworzenie planu analizy danych determinuje efekty, które uzyskamy.

3.1. Zrozumienie uwarunkowań biznesowych i badawczych

a) Określenie celów biznesowych organizacji

Badania postaw i opinii pracowników nie mogą być prowadzone w oderwaniu od kontekstu funkcjonowania organizacji. Powinny być zaprojektowane tak, by były użyteczne w realizowaniu celów organizacji. Należy zatem jednoznacznie określić związek prowadzonych badań z realizacją celów organizacji.

b) Ocena kontekstu prowadzenia analizy

Badania opinii pracowników powinna poprzedzić ocena zagrożeń i kosztów oraz zasobów i możliwości, jakie się z nimi wiążą. Należy określić, czy badanie postaw i opinii jest słusznym przedsięwzięciem na obecnym etapie rozwoju organizacji i czy jego koszty nie przerosną korzyści.

c) Określenie celów analizy danych

W tym zadaniu, w odniesieniu do celów biznesowych organizacji, formułuje się cele analizy danych, określając jakie efekty, przez analizę opinii pracowników, chce się osiągnąć. Znając cel analizy należy zdefiniować zadania analizy danych – sposoby osiągnięcia postawionych celów przez analizę zebranych w badaniu danych. Przykładowo, jeżeli celem badania jest poznanie przyczyn i zmniejszenie rotacji personelu, to za zadanie analizy danych można przyjąć np. budowę modelu predykcyjnego, który pozwalałby przewidywać wartość wskaźnika rotacji na podstawie opinii pracowników.

d) Stworzenie planu analiz danych

Realizacja zadania umożliwi stworzenie planu analizy, uwzględniającej zakładany czas trwania poszczególnych etapów badania, spodziewane efekty zakończenia każdego z nich oraz opis szczegółów technicznych badania. W kontekście badań pracowników ważne jest jasne określenie, w jakiej formie badanie zostanie przeprowadzone (online, w formie papierowej), jakie narzędzia zbierania danych wykorzystamy (kwestionariusz własny czy firmy zewnętrznej) oraz jak zachęcimy pracowników do udziału w badaniu.

3.2. Zrozumienie danych

a) Zebranie wstępnych danych

Realizacja tego zadania odbywa się przez przeprowadzenie zaplanowanego badania właściwego i zgromadzenie danych.

b) Stworzenie opisu danych

Zebrane dane muszą zostać opisane przez dołączenie do każdego zbioru danych definicji poszczególnych zmiennych oraz opisu sposobu zbierania danych.

c) Wstępna eksploracja danych

W tej części analiz należy przeprowadzić wstępny przegląd danych, pozwalający na zauważenie wzorców czy systematycznych błędów pojawiających się w zbiorze danych. Umożliwia to wychwycenie nieprzewidywanych anomalii czy systematycznych braków lub błędów odpowiedzi.

d) Weryfikacja jakości danych

Ostatnim zadaniem etapu zrozumienia danych powinna być analiza reprezentatywności danych oraz określenie ewentualnych grup pracowników, których opinie nie są reprezentowane w badaniu. Ponadto, warto przeprowadzić analizę problemów technicznych pojawiających się podczas zbierania danych i ich konsekwencji, a także przegląd brakujących, niejasnych lub nietypowych odpowiedzi. Zadanie powinno zakończyć się sporządzeniem raportu na temat jakości danych wraz z oszacowaniem ryzyka błędów związanych z używaniem zebranych danych do analiz.

3.3. Przygotowanie danych

a) Wybór danych do analizy

Realizując zadanie należy zdecydować, czy analizie poddany zostanie cały zbiór danych z badania czy wybrane będą jego podzbiory. Trzeba także dokonać wyboru konkretnych zmiennych i przypadków, które zostaną poddane ostatecznej analizie.

b) Czyszczenie danych

W tym zadaniu należy usunąć dane, które uznamy za nadmiarowe, błędne lub niejasne oraz określić sposoby postępowania z obserwacjami nietypowymi i brakami danych.

c) Konstruowanie nowych zmiennych

W tej części procesu analizy wyliczamy i dodajemy nowe zmienne na podstawie zebranych danych, np. sumując wyniki w kwestionariuszach czy obliczając wskaźniki oparte na proporcjach. Dodajemy także zmienne powstałe na skutek przekształceń już istniejących zmiennych.

d) Formatowanie i scalanie danych

Jeśli zachodzi potrzeba, należy scalić dane z innymi. Dane formatujemy i zapisujemy w sposób czytelny dla osób wykonujących ostateczną analizę. Przygotowujemy i sprawdzamy pliki z opisami danych oraz z raportami z wybierania i czyszczenia danych oraz konstruowania nowych zmiennych.

3.4. Modelowanie – analiza danych**a) Wybór techniki modelowania/analizy danych**

Określenie sposobu analizy danych, który pozwoli osiągnąć zakładany cel przez realizację zadania analizy. Istnieje szeroka gama technik analizy, umożliwiających wyjaśnianie, opisywanie, klasyfikowanie czy przewidywanie zjawisk zachodzących w organizacji. Wybór metody analizy odpowiedniej do zakładanego celu może nie być ani łatwy, ani jednoznaczny, ważne zatem, by analityk opisał, jaką analizę wykonuje i dlaczego, tak, by każdy uczestnik procesu badania rozumiał ideę i sens stosowanych analiz.

b) Stworzenie kryteriów testowania wyników

W tym zadaniu następuje ustalenie kryteriów, które powinny spełniać wyniki analiz, by uznać, że osiągnięto postawione zadanie badawcze. Przykładowo, dla modelu predykcyjnego może być to określenie minimalnej, wyjaśnianej przez model zmienności zmiennej zależnej wraz z akceptowalnym poziomem błędu.

c) Budowa modelu/analiza danych

Na tym etapie analitycy z wykorzystaniem wybranych wcześniej metod przetwarzają dane próbując zrealizować zadania analizy danych. W zależności od potrzeb przygotowują określony model, tworzą klasyfikacje czy generują statystyki opisowe.

d) Ocena modelu/wyników analizy

Końcowe wyniki analiz należy ocenić na podstawie ustalonych wcześniej kryteriów. Jeśli wyniki są niezadowolające, należy opisać przyczyny, a następnie powrócić do wcześniejszych etapów analizy, dokonując na nich potrzebnych modyfikacji.

3.5. Ewaluacja**a) Ewaluacja rezultatów**

W tej fazie powinna nastąpić ostateczna ocena wiarygodności i użyteczności rezultatów. Zwykle uzyskiwanych jest wiele różnorodnych wyników analiz danych, należy dokonać wyboru tych, które w największym stopniu odpowiadają postawionym celom.

b) Przegląd procesu analizy

Realizując to zadanie należy opisać dokonane analizy wraz z ich założeniami, informując o tym, co udało się zrealizować dzięki analizie oraz czego nie udało się osiągnąć i dlaczego tak się stało; warto także opisać ograniczenia przeprowadzonych analiz.

c) Wdrożenie działań lub ponowna analiza danych

Na etapie ewaluacji musimy dokonać oceny, czy rezultaty przeprowadzonych analiz dają wiedzę pozwalającą na wdrożenie praktycznych działań. Gdy uznamy, że rezultaty analizy powalają na podjęcie praktycznych działań, powinniśmy zaplanować wdrożenie. Jeśli jednak analiza nie przyniosła rezultatów dających praktyczne zastosowania, należy ponownie rozpocząć cykl analizy danych, modyfikując jej pierwotne założenia.

d) Raport z analizy

Ostatnim zadaniem etapu ewaluacji jest sporządzenie raportu, który oprócz wyników analizy oraz ich opisu, powinien zawierać rekomendacje, co do możliwości prowadzenia dalszych badań postaw i opinii pracowników w danej organizacji.

3.6. Wdrożenie**a) Plan wdrożenia**

Zadanie to polega na oszacowaniu kosztów i zagrożeń, a także możliwości i zysków, jakie niesie wdrożenie pomysłów praktycznego działania wysuniętych na podstawie wyników badania postaw i opinii pracowników. Należy także określić czas, w jakim wdrożenie powinno przynieść określone rezultaty.

b) Plan monitorowania i utrzymania działań

W tym zadaniu, trwającym przez cały czas wdrażania rozwiązań zaproponowanych na podstawie analizy danych, należy monitorować, w jakim stopniu podjęte działania przekładają się na planowane rezultaty. Wyniki nawet najbardziej rzetelnych analiz nie dają gwarancji sukcesu, zatem należy kontrolować efekty ich wdrożenia.

c) Raport końcowy

Ostatnim zadaniem etapu wdrożenia jest, następująca po upływie określonego w planie wdrożenia czasu, ocena, czy działania podjęte na podstawie wyników analiz przyniosły organizacji zakładane korzyści.

4. Podsumowanie

Użyteczność CRISP-DM w badaniach pracowników wynikać może przede wszystkim z faktu, że wprowadza on ład i porządek w proces badań opinii pracowniczych. Model CRISP-DM określa kolejne etapy badania wraz z opisami zadań do wykonania w każdym z nich. Dzięki odwołaniu się do tego formalnego modelu w ciągu całego procesu badania wiadomo, co należy aktualnie zrobić oraz co powinno zostać zrobione w następnym kroku. Powszechnie przyjęty, standardowy model usprawnić może też komunikację pomiędzy uczestnikami procesu badania. Za zasadniczą wadę modelu CRISP-DM można uznać jego raczej opisowy niż wyjaśniający charakter [4]. Model opisuje, co należy zrobić, ale nie precyzuje, w jaki sposób to zrobić, w związku z czym nie może być uznany za szczegółowy algorytm działania, a raczej za zbiór heurystycznych wskazówek. Mimo swej ogólności model CRISP-DM ma przynajmniej cztery zalety, które mogą przemawiać za stosowaniem go w badaniach postaw i opinii pracowników: (1) systematyzuje proces badania i wprowadza ład w proces analizy, (2) akcentuje elastyczność, pokazując, że wracanie do etapów już zakończonych jest możliwe, a często nawet potrzebne, (3) może być użyteczny zarówno dla analityków, jak i dla osób odpowiedzialnych za planowanie i kontrolowanie badań opinii pracowniczych, (4) jest kompleksowy, w ramach modelu CRISP-DM analiza danych rozpoczyna się od zrozumienia potrzeb i problemów organizacji, zatem zanim zgromadzimy jakiegokolwiek dane, natomiast końcowym celem analizy jest podjęcie praktycznych działań.

Bibliografia

1. Alsultanny Y.: Selecting a suitable method of data mining for successful forecasting. *Journal of Targeting, Measurement and Analysis for Marketing*, 19, 3-4, 2011. p. 207-225.
2. Chapman P., Clinton J., Kerber R., Khabaza T., Reinartz T., Shearer C., Wirth R.: *CRISP-DM 1.0 Step-by-step data mining guide*, SPSS Inc. 2000.
3. Christian M.S., Garza A.S., Slaughter J.E.: Work engagement: A quantitative review and test of its relations with task and contextual performance. *Personnel Psychology*, no. 64, 2011, p. 89-136.
4. Cios K.J., Pedrycs W., Swiniarski R.W., Kurgan L.: *Data Mining: A Knowledge Discovery Approach*, Springer, 2011.
5. Davis A.L., Rothstein H.R.: The Effects of the Perceived Behavioral Integrity of Managers on Employee Attitudes: A Meta-analysis. *Journal of Business Ethics*, Vol. 67, no. 4, 2006, p. 407-419.
6. Dirks K.T., Ferrin, D.: The effects of trust in leadership on employee performance, behavior, and attitudes. A meta-analysis, *Academy of Management Best Papers Proceedings Meeting Abstract Supplement 1*, 2000.

7. Hackett R.D.: Work attitudes and employee absenteeism: A synthesis of the literature. *Journal of Occupational Psychology*, Vol. 62, no. 3, 1989, p. 235-248.
8. Harrison D., Newman D., Roth P.L.: How Important Are Job Attitudes? Meta-Analytic Comparisons of Integrative Behavioral Outcomes and Time Sequences. *Academy of Management Journal*, Vol. 49, no. 2, 2006, p. 305-325.
9. Harter J.K., Schmidt F.L., Asplund J.W., Killham E.A., Agrawal S.: Causal Impact of Employee Work Perceptions on the Bottom Line of Organizations, *J.Perspectives on Psychological Science* Vol. 5, no. 4, 2010, p. 378-389.
10. Hiltbrand T.: Behavior-Based Budget Management Using Predictive Analytics. *Business Intelligence Journal*, Vol. 18, no. 1, 2013, p. 25-33.
11. Judge T.A., Kammeyer-Mueller J.D.: Job attitudes. *Annu Rev Psychol*, 63, 2012, p. 341-67.
12. Khaleel M.A.: A Survey of Data Mining Techniques on Medical Data for Finding Locally Frequent Diseases, *International Journal of Advanced Research in Computer Science and Software Engineering*, Vol. 3, no. 12, 2013, p. 149-153.
13. Koys D.: The effects of employee satisfaction, organizational citizenship behavior and Turnover on organizational effectiveness: A unit-level longitudinal study, *Personnel Psychology*, Vol. 54, no. 1, 2001, p. 101-114.
14. Larose D.T.: *Metody i modele eksploracji danych*, PWN, Warszawa 2008.
15. Malhotra N., Mukherjee A.: The relative influence of organisational commitment and job satisfaction on service quality of customer-contact employees in banking call centres, *Journal of Services Marketing*, Vol. 18, no. 3, 2004, p. 162-174.
16. Marbán Ó., Mariscal G., Segovia J.: A Data Mining & Knowledge Discovery Process Model, in: J. Ponce. A. Karahoca (eds.) *Data Mining and Knowledge Discovery in Real Life Applications*, InTech, Rijeka 2009, p. 1-16.
17. McGregor C., Catley C., James A.: A process mining driven framework for clinical guideline improvement in critical care. *CEUR Workshop Proceedings*, 2011, p.765.
18. Oh I.S., Guay R.P., Kim K., Harold C.M., Lee J.H., Heo C.-G., Shin K.-H.: Fit Happens Globally: A Meta-Analytic Comparison of the Relationships of Person-Environment Fit Dimensions with Work Attitudes and Performance Across East Asia, Europe, and North America, *Personnel Psychology*, Vol. 67, no. 1, 2014, p. 99-152.
19. Pereira G.M., Osburn H.G.: Effects of Participation in Decision Making on Performance and Employee Attitudes: A Quality Circles Meta-analysis, *Journal of Business and Psychology*, 22(2), 2007, p. 145-153.
20. Prottas D.J.: Relationships Among Employee Perception of Their Manager's Behavioral Integrity, Moral Distress, and Employee Attitudes and Well-Being, *Journal of Business Ethics*, Vol. 113, no. 1, 2012, p. 51-60.
21. Riedel M., Memon A.S., Memon M.S.: High productivity data processing analytics methods with applications, "Information and Communication Technology, Electronics and Microelectronics (MIPRO), 2014 37th International Convention, 2014, p. 289-294.
22. Riketta M.: The causal relation between job attitudes and performance: a meta-analysis of panel studies. *J Appl Psychol*, Vol. 93, no. 2, 2008, p. 472-81.

23. Rivo E., De La Fuente J., Rivo Á., García-Fontán E., Cañizares M.Á., Gil P.: Cross-Industry Standard Process for data mining is applicable to the lung cancer surgery domain, improving decision making as well as knowledge and quality management. *Clinical and Translational Oncology*, 14, 2012, p. 73–79.
24. Rui-Pingb Z., Tsingan L., Long-Pingb Z.: Role Stressors and Job Attitudes: A Mediated Model of Leader-Member Exchange, *The Journal of social psychology*, Vol. 153, no. 5, 2013, p. 560-576.
25. Shearer C.: The CRIS-DM model: The New Blueprint for Data Mining. *Journal of Data Warehousing* Vol. 5, no. 4, 2000, p. 13–22.
26. Simon D.H, Gómez M.I., McLaughlin E., Wittink D.R.: Employee attitudes, customer satisfaction, and sales performance: assessing the linkages in US grocery stores, *Managerial and Decision Economics*, Vol. 30, no. 1, 2009, p. 27–30.
27. Venter J., de Waal A., Willers C.: Specializing CRISP-DM for evidence mining, [in:] P. Craiger, S. Sheno, S. (eds.) *Advances in Digital Forensic III* p. 303-315, Springer, Boston 2007.
28. Witten I., Frank E.: *Data Mining – Practical Machine Learning Tools and Techniques*, Elsevier, 2005.
29. Xiao X., Xu H., Xu S.: Using IBM SPSS modeler to improve undergraduate mathematical modelling competence. *Computer Applications in Engineering Education*, in Press, 2015.

Abstract

Cross Industry Standard Process for Data Mining (CRISP–DM) is a model which describes data analysis as a coherent process and distinguishes appropriate stages to be carried out. These stages are as following: (1) business understanding, (2) data understanding, (3) data preparation, (4) modelling/data analysis, (5) evaluation of results, (6) deployment. Previous studies have demonstrated effectiveness of CRISP – DM as a data analysis strategy, hence the idea that using CRISP – DM as a model of conducting employees attitudes and opinion research can be beneficial. For first, CRISP – DM is a set of guidelines and best practises which improve research management and knowledge discovery from collected data. For second, one standard model will facilitate communication between specialists on different levels of organization structures. Interest in employee opinion research is growing in contemporary business organisation, so it seems to be important to put forward a systematic model of conducting this kind of research.