

## Wykorzystanie modelu CART-Logit do analizy fałszerstw sprawozdań finansowych

Marek Sylwestrzak\*

**Streszczenie:** *Cel* – Opracowanie modelu hybrydowego CART-Logit wykrywającego fałszerstwa sprawozdań finansowych na podstawie danych finansowych amerykańskich spółek giełdowych oskarżonych z przepisu 10(b)-5 ustawy Securities Exchange Act w latach 2000–2007.

*Metodologia badania* – W opracowaniu zastosowano model hybrydowy CART-Logit z wykorzystaniem 10 zmiennych finansowych.

*Wynik* – Przeprowadzone analizy wykazały, że model hybrydowy wykazał większą zdolność predykcyjną niż zwykła regresja logistyczna. Uwzględnienie w modelu wskaźnika predykcji bankructwa Altmana zwiększyło trafność metody. Uzyskane wyniki potwierdzają, że najważniejszą pozycją przy wykrywaniu fałszerstw sprawozdań finansowych są środki pieniężne.

*Oryginalność/wartość* – Artykuł stanowi empiryczną analizę możliwości wykrycia fałszerstw sprawozdań finansowych za pomocą nowej metody badawczej.

**Słowa kluczowe:** regresja logistyczna, drzewa decyzyjne, oszustwa księgowe, rynek amerykański

### Wprowadzenie

Rozmiary fałszowania wyników finansowych przedsiębiorstw od długiego czasu budzą zainteresowanie analityków, regulatorów, badaczy i inwestorów (Beneish, 1999). Fałszowanie danych finansowych prowadzi do złamania nadrzędnej w rachunkowości zasady prawdziwego i wiernego obrazu oraz stwarza duże zagrożenie dla otoczenia, a w szczególności inwestorów, wpływając na podejmowanie błędnych decyzji inwestycyjnych (Ata, Seyrek, 2009). Ponadto jest to zjawisko niebezpieczne, gdyż wiąże się z ukrywaniem prawdziwej sytuacji spółki, zwłaszcza możliwości jej bankructwa.

Jak zauważają Summers i Sweeney (1998), przyczyną fałszerstw sprawozdań finansowych jest dążenie do podwyższenia ceny akcji spółki, ponieważ dzięki manipulacjom finansowym członkowie jej zarządu mogą korzystnie sprzedać posiadane przez siebie akcje spółki. Ponadto członkowie zarządu starają się spełnić oczekiwania inwestorów co do wyników finansowych spółki prognozowanych przez analityków finansowych. Według Basu (1997) „złe wiadomości” znajdują odzwierciedlenie w zmianie ceny akcji znacznie

\* mgr Marek Sylwestrzak, Uniwersytet Warszawski, Wydział Nauk Ekonomicznych, e-mail: msylwestrzak@wne.uw.edu.pl.

szybciej niż „dobre wiadomości”. Givoly i Hayn (2002) stwierdzili zaś, że w ostatnich latach inwestorzy coraz szybciej reagują na „złe wiadomości” i dlatego są one rzadziej ujawniane przez kierownictwa przedsiębiorstw lub następuje to z pewnym opóźnieniem, zwłaszcza gdy pojawiają się one dopiero podczas ogłaszania okresowych wyników spółki.

W artykule za definicję fałszerstwa sprawozdania finansowego przyjęto zapis z przepisu 10(b)-5 ustawy *Securities Exchange Act* z 1933 roku, według którego niezgodne z prawem są wszelkie działania, które bezpośrednio lub pośrednio wykorzystują wszelkie środki w celu oszukania, składania fałszywych zeznań, pominięcia istotnych informacji lub w inny sposób prowadzą do ujęcia operacji gospodarczych mogących zmylić inną osobę przy przeprowadzaniu transakcji akcjami i innymi papierami wartościowymi.

Celem pracy jest opracowanie modelu hybrydowego CART-Logit na podstawie danych ze spółek notowanych na amerykańskim rynku kapitałowym, który trafnie klasyfikowałby sprawozdania finansowe jako sfalszowane. Dotychczasowe analizy, związane z wykrywaniem fałszerstw sprawozdań finansowych, biorą pod uwagę zarówno okresy koniunktury, jak i dekonunktury, co może wpłynąć na uzyskane wyniki. W przedstawionych badaniach wykorzystuje się ponadto regresje logistyczne lub metody Data Mining, zwłaszcza drzew decyzyjnych, ale bez próby połączenia estymacji obydwu modeli.

## 1. Przegląd literatury

Persons (1995) przeprowadził analizy, które wykazały, że istotnymi zmiennymi mogącymi pomóc przy wykrywaniu fałszerstw jest wskaźnik ogólnego zadłużenia, wskaźnik rotacji aktywów ogółem, udział aktywów obrotowych w aktywach ogółem oraz logarytm wielkości aktywów ogółem.

Summers i Sweeney (1998) stwierdzili, że spółki fałszujące dane finansowe miały wyższy wskaźnik rotacji zapasów, rentowności aktywów ogółem i większy wzrost przychodów ze sprzedaży w ciągu trzech lat przed wystąpieniem oszustwa.

Fich i Shivdasani (2007) wykazali, że słabe struktury zarządcze powodują zwiększenie prawdopodobieństwa fałszowania danych finansowych. Wielkość firmy zmniejszała natomiast prawdopodobieństwo zastosowania kreatywnych technik księgowych.

Johnson i inni (2009) wśród 20 wskaźników wskazali cztery zmienne, które dodatkowo wpływały na wzrost wykrycia fałszowania sprawozdań finansowych: przyznanie członkom zarządu akcji spółki, które mogą zostać sprzedane w każdym momencie, wzrost wielkości sprzedaży w ciągu trzech lat przed dokonaniem oszustwa, zasiadanie przez *insiders* w Komitecie audytu oraz wielkość logarytmu liczby spotkań zarządu w ciągu roku.

Dechow i inni (2011) wykazali, że firmy fałszujące sprawozdania finansowe miały wyższy wskaźnik rotacji zapasów, rotacji należności, poziomu rozliczeń międzyokresowych, udziału aktywów obrotowych bez środków pieniężnych w aktywach ogółem, zmianę przychodów ze sprzedaży oraz miały niższy wskaźnik rentowności aktywów ogółem.

Spathis (2002) wśród 10 analizowanych zmiennych wskazał pięć wskaźników, które są pomocne przy wykrywaniu kreatywnych technik: rotację zapasów, rentowność aktywów ogółem, wskaźnik zadłużenia ogółem, stosunek kapitału pracującego do aktywów ogółem i wskaźnik predykcji bankructwa Altmana, przy czym uwzględnienie wskaźnika Altmana zwiększyło trafność modelu o 1,2%.

Analiza przeprowadzona przez Gupta i Gilla (2012) określiła siedem wskaźników, istotnych przy wykrywaniu fałszowania sprawozdań finansowych odwołujących się do wielkości środków pieniężnych, rentowności aktywów ogółem, logarytmu aktywów ogółem, stosunku środków pieniężnych do aktywów obrotowych, zapasów do zobowiązań krótkoterminowych, stosunku – środków pieniężnych, kapitałów własnych, zapasów i zysku zatrzymanego do aktywów ogółem.

Inne wskaźniki pomocne przy analizie zjawiska fałszowania sprawozdań zostały wskazane przez Kotsiantisa i in. (2006). Były to: ogólne zadłużenie, zadłużenia długoterminowe, stosunek należności do aktywów ogółem, kapitału pracującego do aktywów ogółem, środków pieniężnych do aktywów obrotowych, aktywów trwałych do aktywów ogółem i cyklu zobowiązań w dniach. Jak wskazali autorzy stosunkowo niewielka liczba wskaźników finansowych w dużej mierze determinuje wyniki klasyfikacji modelu.

Z kolei Ata i Seyrek (2009) wykazali, że wskaźnikami służącymi do wykrywania fałszowań sprawozdań finansowych są: wskaźnik ogólnego zadłużenia, rentowność aktywów ogółem, udział zapasów w aktywach ogółem i rentowność sprzedaży brutto.

Badanie przeprowadzone przez Pai i in. (2011) pozwoliło natomiast na określenie sześciu wskaźników istotnych przy detekcji manipulacji danymi finansowymi: stosunek zysku netto do aktywów trwałych, rentowność aktywów ogółem, wielkość EBIT, rotację zapasów, wskaźnik ogólnego zadłużenia oraz liczbę posiadanych akcji przez członków rady dyrektorów, przy czym według autorów najważniejszym wskaźnikiem była ostatnia zmienna.

## 2. Metodologia badania

W pracy, na podstawie przedstawionej w poprzednim rozdziale literatury, przyjęto następującą postać estymowanego modelu CART-Logit:

$$fraud_i = \beta_0 + \beta * X_i + \varepsilon_i,$$

gdzie:

$fraud_i$  – zmienna zero-jedynkowa przyjmującą wartość 1, jeżeli spółka została oskarżona przez SEC z przepisu 10(b)-5, a zero – w przeciwnym przypadku,

$\beta_0$  – wyraz wolny modelu,

$X_i$  – macierz zmiennych objaśniających,

$\varepsilon_i$  – błąd losowy.

Na podstawie analizy zmiennych pod kątem liczby ich wystąpień w pracach badawczych, na podstawie ich istotności wybrano 12 zmiennych finansowych oraz wskaźnik Altmana na potrzeby weryfikacji drugiej hipotezy badawczej (tab. 1).

**Tabela 1**

Wybór zmiennych do modelu

Zmienna	Formuła	Liczba zastosowań
CASH_PART_CURRENT	środki pieniężne/aktywa obrotowe	2
CURRENT_RATIO	aktywa bieżące/zobowiązania bieżące	4
DEBT_RATIO	zobowiązania ogółem/aktywa ogółem	7
INVENTORY_PART	zapasy/aktywa ogółem	6
LOG_AT	logarytm naturalny aktywów ogółem	3
LTD_PART	zobowiązania długoterminowe/aktywa ogółem	3
NET_PROFIT_RATIO	zysk netto/przychody netto ze sprzedaży	4
QUICK_RATIO	(aktywa bieżące – zapasy)/zobowiązania bieżące	3
RECEIVABLE_PART	należności/aktywa ogółem	5
RETAINED_RATIO	zysk zatrzymany/aktywa ogółem	3
ROA	zysk netto/aktywa ogółem	5
TAT_RATIO	przychody netto ze sprzedaży/aktywa ogółem	5
ALTMAN_SCORE	wskaźnik Altmana z 1968 r.	5

Źródło: opracowanie własne na podstawie analizowanych artykułów.

Badaniem objęto grupę amerykańskich spółek giełdowych, dobranych na podstawie pozwów SEC zawartych w bazie AAER w latach 2000–2013. Najpierw stworzono listę spółek, które w latach 2000–2007 dokonały fałszerstw sprawozdań finansowych. Analizowany okres zawiera się między załamaniem bańki internetowej na rynku amerykańskim w połowie 2000 roku a początkiem kryzysu finansowego w 2007 roku w celu pominięcia w analizie czynnika makroekonomicznego, który mógłby zaburzyć uzyskane wyniki. Następnie z wytypowanej grupy wykluczono spółki:

- nieoskarżone o manipulacje z przepisu 10(b)-5,
- należące do sektora finansowego na podstawie kodu SIC dostępnego w bazie EDGAR SEC,
- których danych finansowych nie było w bazie EDGAR SEC.

Ostatecznie grupa badawcza objęła 181 spółek, co oznacza 465 obserwacji rocznych. Następnie zostały dobrane spółki niefałszujące danych finansowych na podstawie:

- przynależności do tego samego sektora gospodarki według czterocyfrowego kodu SIC,
- wielkości, określonej na podstawie wartości aktywów ogółem, minimalizującej różnicę pomiędzy spółkami, jeżeli w danym sektorze gospodarki nie znaleziono odpowiedniej spółki, powtarzano pierwszy krok, dobierając spółkę z trzycyfrowego, a następnie dwucyfrowego kodu SIC,
- dostępności danych w tym samych okresach, co spółki fałszujące dane.

Na podstawie przedstawionych kryteriów grupę kontrolną tworzyło 327 spółek, tj. 920 obserwacji rocznych. Łącznie na badaną próbę składało się 508 przedsiębiorstw, tj. 1385 obserwacji rocznych.

W analizowanych artykułach badacze, stosując różne modele, dążyli do osiągnięcia maksymalnej trafności proponowanych modeli diagnostycznych. Jednakże w żadnym nie uwzględniono podejścia hybrydowego, tj. połączenia regresji liniowej ze zmiennymi z drzew decyzyjnych, które cechuje wyższa trafność predykcji (Łapczyński, 2014). W tym przypadku pierwsza hipoteza badawcza została sformułowana następująco:

**H<sub>1</sub>: Model hybrydowy CART-Logit zwiększy trafność prognostyczną analizy w porównaniu z wynikami regresji logistycznej.**

W części badań autorzy uwzględniali model predykcji bankructwa Altmana z 1968 roku, co zwiększało dokładność prognostyczną modelu. W związku z tym postanowiono zweryfikować następującą hipotezę:

**H<sub>2</sub>: Zastosowanie modelu Altmana (1968) zwiększa trafność dopasowania modelu hybrydowego.**

W żadnym z przeanalizowanych artykułów nie dokonano analizy fałszerstw sprawozdań finansowych tylko jedną techniką. Spółki stosujące te metody mogą zarówno zwiększać albo zmniejszać dane finansowe, co może powodować zmniejszenie dokładności proponowanych modeli. W celu weryfikacji tego stwierdzenia została postawiona następująca hipoteza pomocnicza:

**H<sub>3</sub>: Rozszerzenie modelu hybrydowego o rodzaj fałszerstw sprawozdań finansowych zwiększy trafność proponowanej metody badawczej.**

### 3. Estymacja modelu

Przed przystąpieniem do badania przeprowadzono test Manna-Whitney'a na równość tendencji centralnych pomiędzy spółkami z grupy kreatywnej i kontrolnej. Test wskazał, że dla zmiennej rentowności aktywów ogółem (**ROA**) oraz stosunku przychodów ze sprzedaży do aktywów ogółem (**TAT\_RATIO**) nie można było odrzucić hipotezy zerowej testu na poziomie istotności 10%, dlatego nie zostały one uwzględnione w analizie. Ponadto przeprowadzono test VIF (*Variance Inflation Factor*) na występowanie współliniowości, który wskazał, że w modelu występuje współliniowość, gdyż dla zmiennych **QUICK\_RATIO** i **CURRENT\_RATIO** wartości były powyżej 10. Zdecydowano więc o usunięciu pierwszego wskaźnika ze względu na wyższą wartość w teście VIF, co po powtórzeniu testu wyeliminowało współliniowość w modelu.

W przypadku zastosowania 9 zmiennych model drzewa decyzyjnego wskazał 8 wskaźników (bez **LTD\_PART**) będących jego węzłami, przy czym podstawę drzewa stworzył wskaźnik udziału środków pieniężnych w aktywach obrotowych (**CASH\_PART\_CURRENT**), a węzłami pierwszego poziomu były wskaźniki logarytmu aktywów ogółem (**LOG\_AT**) oraz wskaźnik zysku zatrzymanego do aktywów ogółem (**RETAINED\_RATIO**). Trafność drzewa wyniosła – 75,2%, wrażliwość – 36,1%, a specyficzność – 95,0%.

Na podstawie predykcji drzewa zostały stworzone trzy zmienne zero-jedynkowe, które zostaną dołączone do estymacji modelu logitowego:

- **LISC1** – przyjmujący wartość jeden, jeżeli wskaźnik **CASH\_PART\_CURRENT** jest mniejszy od 0,2134, a **LOG\_AT** jest większy bądź równy 8,171,
- **LISC2** – przyjmujący wartość jeden, jeżeli wskaźnik **CASH\_PART\_CURRENT** jest mniejszy od 0,2134, a **LOG\_AT** jest mniejszy od 8,171,
- **LISC3** – przyjmujący wartość jeden, jeżeli wskaźnik **CASH\_PART\_CURRENT** jest większy bądź równy 0,2134, a **RETAINED\_RATIO** jest mniejszy od 0,07044.

Regresja logistyczna dla 9 zmiennych (model 1) wykazała, że trzy zmienne są istotne przy wykrywaniu fałszerstw sprawozdań finansowych. Przeprowadzenie analizy dla modelu z ograniczeniami (model 2) wykazało, że oprócz – logarytmu wielkości aktywów – pozostałe dwie zmienne są istotne. Wartość uzyskanej statystyki testu ilorazu wiarygodności (*likelihood ratio test* – LRT) na poziomie 0,28 oznacza, że należy wybrać model z ograniczeniami. Również wartość testu Hosmera-Lemeshowa na poziomie 0,48 wskazuje na prawidłowe dopasowanie formy funkcyjnej modelu z ograniczeniami.

Wzrost wartości wskaźnika udziału środków pieniężnych w aktywach obrotowych wpływa na zmniejszenie prawdopodobieństwa zakwalifikowania spółki jako stosującej kreatywne techniki, natomiast wzrost wartości aktywów ogółem oraz poziomu zysków zatrzymanych w sumie bilansowej zwiększa takie prawdopodobieństwo.

Następnie zbadano trafność, wrażliwość i specyficzność modelu, posługując się tabelą trafień i zmieniając punkt graniczny z 0,5 na 0,33, gdyż model nie jest zbilansowany, czyli liczba obserwacji spółek fałszujących dane finansowe do liczby firm niestosujących kreatywnych technik wynosi 1:2. Zgodnie z wynikami, trafność wyniosła 54,5%, wrażliwość 63,2%, a specyficzność 50,1%.

W celu weryfikacji pierwszej hipotezy w regresji logistycznej z ograniczeniami zostały uwzględnione trzy zmienne stworzone na podstawie modelu drzewa decyzyjnego (model 3). Istotne okazały się zmienne **LISC2** i **LISC3**, więc w kolejnej regresji nie uwzględniono zmiennej **LISC1** (model 4). Dwa wskaźniki z modelu drzewa decyzyjnego zwiększały prawdopodobieństwo uznania spółki za kreatywną, a także wzrosły wszystkie trzy parametry z tabeli trafień, w związku z tym nie ma podstaw do odrzucenia pierwszej hipotezy.

W celu weryfikacji drugiej hipotezy w modelu z ograniczeniami uwzględniono wskaźnik predykcji bankructwa Altmana (model 5). Zmienna okazała się statystycznie istotna, a wraz ze wzrostem jej wartości maleje prawdopodobieństwo zakwalifikowania spółki jako kreatywnej. Trafność modelu ze wskaźnikiem Altmana była wyższa o 4,9 punktu procentowego, a wrażliwość i specyficzność odpowiednio o 1,1 oraz 6,7 punktu procentowego, co oznacza, że nie ma podstaw do odrzucenia drugiej hipotezy.

W celu weryfikacji trzeciej hipotezy badawczej zostały utworzone dodatkowo dwie zmienne zależne:

- **fsf1** – zmienna zero-jedynkowa przyjmująca wartość 1, jeżeli spółka zawyżała wielkość przychodów,
- **fsf2** – zmienna zero-jedynkowa przyjmująca wartość 1, jeżeli spółka pomniejszała wielkość swoich zobowiązań.

**Tabela 2**

Wyniki regresji dla poszczególnych modeli

Model	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Zmienna zależna	fraud	fraud	fraud	fraud	fraud	fsf1	fsf2
CASH_PART_CURRENT	-1,38*** (0,00)	-1,17** (0,01)	-0,55 (0,12)	-0,81*** (0,01)	-0,90*** (0,00)	-2,45*** (0,00)	-1,68*** (0,00)
CURRENT_RATIO	0,01 (0,42)						
DEBT_RATIO	-0,56 (0,19)						
INVENTORY_PART	-0,55 (0,15)						
LOG_AT	0,06* (0,09)	0,05 (0,12)	0,09** (0,01)	0,11*** (0,00)	0,11** (0,00)	0,05 (0,27)	-0,04 (0,49)
LTD_PART	0,62 (0,24)						
NET_PROFIT_RATIO	-0,28 (0,56)						
RECEIVABLE_PART	0,53 (0,38)						
RETAINED_RATIO	-0,08** (0,02)	0,07** (0,05)	-0,05 (0,17)	-0,06* (0,10)	0,04 (0,54)	0,03 (0,59)	0,16 (0,24)
LISC1			0,41 (0,14)				
LISC2			0,89*** (0,00)	0,70*** (0,00)	0,70*** (0,00)	0,09 (0,70)	0,14 (0,64)
LISC3			0,53*** (0,00)	0,47*** (0,01)	0,43** (0,01)	0,73*** (0,00)	0,66** (0,03)
ALTMAN_SCORE					-0,09* (0,06)		
Stała	-0,44 (0,20)	-0,57** (0,01)	-1,60*** (0,00)	-1,50*** (0,00)	-1,30*** (0,00)	-0,51 (0,21)	-0,10 (0,86)
Liczba obserwacji	1 385	1 385	1 385	1 385	1 385	830	573
Trafność (%)	56,0	54,5	60,0	58,6	59,4	58,0	57,1
Wrażliwość (%)	63,7	63,2	62,6	63,7	64,3	70,9	62,9
Specyficzność (%)	52,1	50,1	58,7	56,1	56,8	51,8	54,4

\*\*\*  $p$ -value < 0,01; \*\*  $p$ -value < 0,05; \*  $p$ -value < 0,10.

Źródło: opracowanie własne w programie RStudio.

Dla żadnego z dwóch modeli nie uzyskano wyższej trafności niż dla standardowego modelu hybrydowego, w związku z tym należy odrzucić trzecią hipotezę. Warto jednak zaznaczyć, że w obydwu modelach tylko dwie zmienne były istotne, co świadczy, że w analizach należałoby uwzględnić technikę fałszowania sprawozdań finansowych.

## Uwagi końcowe

Fałszowanie sprawozdań finansowych, choć nie jest najczęściej występującym rodzajem oszustwa finansowego, generuje jednak najwyższą stratę (ACFE, 2014). Przedstawiona literatura, dotycząca analizowanego zjawiska, wykorzystuje zarówno regresje logistyczne, jak i drzewa decyzyjne, przy czym uwzględnia zarówno lata koniunktury i dekonunktury, co może wpływać na uzyskane wyniki.

Przeprowadzone analizy wykazały, iż nie można odrzucić hipotezy, że model hybrydowy osiągnie wyższą trafność prognostyczną. Potwierdzono także drugą hipotezę, iż uwzględnienie modelu Altmana zwiększa zdolność predykcyjną modelu. Należałoby natomiast odrzucić trzecią hipotezę, że uwzględnienie rodzaju fałszerstw poprawi trafność regresji, choć w tym wypadku wpływ na wyniki mogły mieć inne wartości graniczne dla zmiennych z drzewa decyzyjnego.

Wyniki badań wskazują, że do klasyfikacji spółek wystarczy mała liczba zmiennych, co jest zgodne z postulatami wysuwanymi przez Grove'a i Cooka (2004) oraz Zacka (2012), według których przy weryfikacji sprawozdań finansowych należy uwzględnić tylko niektóre wskaźniki finansowe, których wartości mogą wskazywać na istnienie nieprawidłowości w spółce. Przedstawione badanie skupia się na rynku amerykańskim, ponieważ w Stanach Zjednoczonych istnieje odpowiednia baza, dzięki której można jednoznacznie wskazać spółki fałszujące dane finansowe. Jednak należy podkreślić, że polscy badacze również prowadzą prace empiryczne dotyczące obszaru manipulowania wynikami finansowymi przez przedsiębiorstwa (Wójtowicz, 2010, s. 129–169; Piosik, 2016, s. 103–282).

Nie można natomiast zbudować jednego zestawu standardów rachunkowości, które ograniczyłyby istnienie kreatywnej księgowości, gdyż – jak zaznacza Sunder (2010) – stworzenie jednego zestawu jest niewykonalne i trudno byłoby przewidzieć, które rozwiązania powinny być dopuszczone. Ograniczenie możliwości alternatywnych rozwiązań i ujednolicenie standardów prowadziłyby do zmniejszenia znaczenia zawodu księgowego.

## Literatura

- Association of Certified Fraud Examiners (2014). *Report to the Nations on Occupational Fraud and Abuse. 2014 Global Fraud Study*.
- Ata, H., Seyrek, I. (2009). The Use of Data Mining Techniques in Detecting Fraudulent Financial Statements: An Application on Manufacturing Firms. *The Journal of Faculty of Economics and Administrative Sciences*, 14 (2), 157–170.



- Basu, S. (1997). The Conservatism Principle and the Asymmetric Timeliness of Earnings. *Journal of Accounting and Economics*, 24 (1), 3–37.
- Beasley, M. (1996). An Empirical Analysis of the Relation between the Board of Director Composition and Financial Statement Fraud. *Accounting Review*, 71 (4), 443–465.
- Beneish, M. (1999). The Detection of Earnings Manipulation. *Financial Analysts Journal*, 55 (5), 24–36.
- Dechow, P., Ge, W., Larson, C., Sloan, R. (2011). Predicting Material Accounting Misstatements. *Contemporary Accounting Research*, 28 (1), 17–82.
- Fich, E.M., Shivdasani, A. (2007). Financial Fraud, Director Reputation, and Shareholder Wealth. *Journal of Financial Economics*, 86 (2), 306–333.
- Givoly, D., Hayn, C. (2002). Rising Conservatism: Implications for Financial Analysis. *Financial Analysts Journal*, 58 (1), 56–74.
- Grove, H., Cook, T. (2004). Lessons for Auditors: Quantitative and Qualitative Red Flags. *Journal of Forensic Accounting*, 5 (1), 131–146.
- Gupta, R., Gill, N. (2012). Prevention and Detection of Financial Statement Fraud – An Implementation of Data Mining Framework. *Editorial Preface*, 3 (8), 150–160.
- Johnson, S., Ryan, H., Tian, Y. (2009). Managerial Incentives and Corporate Fraud: The Sources of Incentives Matter. *Review of Finance*, 13 (1), 115–145.
- Kotsiantis, S., Koumanakos, E., Tzelepis, D., Tampakas, V. (2006). Forecasting Fraudulent Financial Statements Using Data Mining. *International Journal of Computational Intelligence*, 3 (2), 104–110.
- Loh, W. (2011). Classification and Regression Trees. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 1 (1), 14–23.
- Łapczyński, M. (2014). Modele hybrydowe C&RT-Logit w analizie migracji klientów. *Studia Ekonomiczne*, 195, 90–102.
- Pai, P., Hsu, M., Wang, M. (2011). A Support Vector Machine-Based Model for Detecting Top Management Fraud. *Knowledge-Based Systems*, 24 (2), 314–321.
- Persons, O. (1995). Using Financial Statement Data to Identify Factors Associated with Fraudulent Financial Reporting. *Journal of Applied Business Research*, 11 (3), 38–46.
- Piosik, A. (2016). *Kształtowanie wyniku finansowego przez podmioty sprawozdawcze w Polsce. Diagnoza dobrej i złej praktyki w rachunkowości*. Katowice: Wydawnictwo Uniwersytetu Ekonomicznego w Katowicach.
- Rezaee, Z. (2005). Causes, Consequences, and Deterrence of Financial Statement Fraud. *Critical Perspectives on Accounting*, 16 (3), 277–298.
- Spathis, C. (2002). Detecting False Financial Statements Using Published Data: Some Evidence From Greece. *Managerial Auditing Journal*, 17 (4), 179–191.
- Sunder, S. (2010). Adverse Effects of Uniform Written Reporting Standards on Accounting Practice, Education, and Research. *Journal of Accounting and Public Policy*, 29 (2), 99–114.
- Summers, S., Sweeney, J. (1998). Fraudulently Misstated Financial Statements and Insider Trading: An Empirical Analysis. *Accounting Review*, 73 (1), 131–146.
- Wójtowicz, P. (2010). *Wiarygodność sprawozdań finansowych wobec aktywnego kształtowania wyniku finansowego*. Kraków: Wydawnictwo Uniwersytetu Ekonomicznego w Krakowie.
- Zack, G. (2012). *Financial Statement Fraud: Strategies for Detection and Investigation*. New Jersey: John Wiley & Sons.

### THE HYBRID CART-LOGIT MODEL APPLICATION IN THE DETECTION OF FALSIFIED FINANCIAL STATEMENT

**Abstract:** *Purpose* – The elaboration a hybrid CART-Logit model to detection of the financial statement fraud based on the financial data from the US companies accused by the US Securities and Exchange Commission manipulating financial statements of the rule 10(b)-5 Securities Exchange Act between 2000–2007.

*Design/methodology/approach* – In the study a hybrid CART-Logit model is used with ten financial ratios.

*Findings* – The results confirm that a hybrid model had greater predictive power than ordinary logistic regression. The inclusion of the Altman model increased the accuracy of the method. The analysis confirmed that the most sensitive position in financial statement is cash.

*Originality/value* – The article is an empirical analysis of capabilities in detection of financial statements fraud based on new research method.

**Keywords:** logit regression, decision trees, accounting fraud, American market

#### Cytowanie

Sylwestrzak, M. (2017). Wykorzystanie modelu CART-Logit do analizy fałszerstw sprawozdań finansowych. *Finanse, Rynki Finansowe, Ubezpieczenia*, 4 (88/1), 403–412. DOI: 10.18276/frfu.2017.88/1-39.