

Metody oceny jakości modeli prognozowania kondycji ekonomiczno-finansowej przedsiębiorstw

Zbigniew Omiotek¹

1. Wstęp

Ocena „jakości” modelu polega na porównaniu wyników generowanych przez model z informacją o rzeczywistym stanie przedsiębiorstwa. W procesie oceny uwzględnia się dwa kryteria:

- sprawność modelu,
- kalibrację modelu.

Sprawność określa, jak dobrze model rozdziela dwie grupy przedsiębiorstw na bankrutów i jednostki niezagrożone upadkiem [6]. W procesie oceny sprawności istotne jest, czy model prawidłowo klasyfikuje obiekty do odpowiedniej grupy. Natomiast nie jest istotne, z jakim prawdopodobieństwem obiekty są klasyfikowane do grup.

Kalibracja określa, z jaką siłą oszacowane prawdopodobieństwa zgadzają się ze stanem rzeczywistym. Decydującą rolę odgrywają tutaj wielkości oszacowanych prawdopodobieństw – im są one wyższe tym lepszy jest oceniany model.

Dokonując analizy porównawczej, należy spełnić 2 warunki:

1. Można porównywać wyniki poszczególnych modeli uzyskane podczas badania tej samej próby jednostek.
2. Obiekty z próby badawczej nie mogą być wcześniej wykorzystywane do budowy modeli.

Wśród wielu metod oceny jakości modeli, do najbardziej popularnych należą:

w zakresie oceny sprawności

- macierz klasyfikacji przedsiębiorstw,
- iloraz szans,
- krzywe koncentracji:
 - o **CAP (Cumulative Accuracy Profiles)**,
 - o **ROC (Relative or Receiver Operating Characteristic)**,
- wskaźnik dokładności (Giniego),

w zakresie oceny kalibracji

- wskaźnik Briera,
- wskaźnik wiarygodności modelu prognozowania zagrożenia finansowego przedsiębiorstw – L (model) (*likelihood of the model*) [4].

Znane są również inne metody stosowane w analizie porównawczej modeli:

- R-kwadrat McFaddena (pseudo R-kwadrat),
- R-kwadrat Efrona,
- R-kwadrat Cragga-Uhlera,
- R-kwadrat Buse'a,
- Miara Aldricha-Nelsona,

- Miara Vealla-Zimmermanna,
- Nowa miara Estreli,
- Skorygowany zliczeniowy R-kwadrat,
- R-kwadrat Ben-Akivy i Lermana,
- Miara λ Cramera.

Jednak największą popularnością i uznaniem cieszy się w praktyce sześć metod wymienionych wcześniej [3].

2. Metody oceny sprawności modeli

2.1. Macierz klasyfikacji przedsiębiorstw

Do budowy macierzy klasyfikacji wykorzystane są następujące parametry:

P₁ – liczba bankrutów zaklasyfikowana do grupy przedsiębiorstw upadłych.

NP₁ – liczba bankrutów zaklasyfikowana do grupy przedsiębiorstw nieupadłych.

P₂ – liczba przedsiębiorstw nieupadłych zaklasyfikowana do grupy przedsiębiorstw nieupadłych.

NP₂ – liczba jednostek nieupadłych zaklasyfikowana do grupy bankrutów.

Zdefiniowane wyżej parametry służą następnie do obliczenia wskaźników oceny jakości modeli.

Sprawność I rodzaju (SP₁) – określa jaki procent bankrutów został prawidłowo sklasyfikowany.

$$(1) \quad SP_1 = \frac{P_1}{P_1 + NP_1} \cdot 100\%$$

Błąd I rodzaju (B₁) – określa jaki procent bankrutów został nieprawidłowo sklasyfikowany.

$$(2) \quad B_1 = \frac{NP_1}{P_1 + NP_1} \cdot 100\%$$

Sprawność II rodzaju (SP₂) – określa jaki procent przedsiębiorstw niezagrażonych upadkiem został prawidłowo sklasyfikowany.

$$(3) \quad SP_2 = \frac{P_2}{P_2 + NP_2} \cdot 100\%$$

Błąd II rodzaju (B₂) – określa jaki procent przedsiębiorstw niezagrażonych upadkiem został nieprawidłowo sklasyfikowany.

$$(4) \quad B_2 = \frac{NP_2}{P_2 + NP_2} \cdot 100\%$$

Sprawność ogólna (SP) – określa jaki procent wszystkich przedsiębiorstw został prawidłowo sklasyfikowany.

$$(5) \quad SP = \frac{P_1 + P_2}{P_1 + NP_1 + P_2 + NP_2} \cdot 100\%$$

Błąd ogólny (B) – określa jaki procent wszystkich przedsiębiorstw został nieprawidłowo sklasyfikowany.

$$(6) \quad B = \frac{NP_1 + NP_2}{P_1 + NP_1 + P_2 + NP_2} \cdot 100\%$$

¹ Wyższa Szkoła Zarządzania i Administracji, Katedra Informatyki i Inżynierii Wiedzy.

Macierz klasyfikacji jest odpowiednią metodą oceny modeli prognozowania zagrożenia finansowego przedsiębiorstw, pod warunkiem, że mniej istotna jest z punktu przeprowadzonej analizy ocena ryzyka działalności przedsiębiorstwa.

2.2. Iloraz szans

Iloraz szans oferuje podobne możliwości oceny modeli prognozowania zagrożenia finansowego przedsiębiorstw, co macierz klasyfikacji. Sprawność określana jest za pomocą stosunku iloczynu liczb poprawnie zaklasyfikowanych przedsiębiorstw do iloczynu liczb nieprawidłowo zaklasyfikowanych jednostek.

Iloraz szans liczony jest według następującej zależności:

$$(7) \quad IS = \frac{P_1 \cdot P_2}{NP_1 \cdot NP_2}$$

gdzie: IS – iloraz szans,

P_1, P_2, NP_1, NP_2 – oznaczenia, jak w przypadku macierzy klasyfikacji.

Im wyższa jest wartość ilorazu szans, tym lepszy jest oceniany model. Iloraz szans posiada podobne ograniczenia w zakresie oceny modeli, co macierz klasyfikacji.

2.3. Krzywe koncentracji

Krzywe koncentracji pozwalają na graficzne porównanie wyników klasyfikacji przedsiębiorstw uzyskanych za pomocą wielu modeli. Stosowane są dwa rodzaje krzywych koncentracji – CAP oraz ROC. Obydwie krzywe dostarczają takich samych wyników, lecz prezentują je w nieco inny sposób.

Podczas budowy krzywych koncentracji CAP zachowana jest następująca metodyka postępowania [2, 5]:

1. Przedsiębiorstwa są porządkowane według wartości funkcji modelu – od najgorszych do najlepszych wyników.
2. Wyznaczony zostaje procentowy skumulowany udział przedsiębiorstw w próbie dla każdej obserwacji.
3. Następuje przyporządkowanie poszczególnym wartościom procentowego skumulowanego udziału przedsiębiorstw z całej próby procentowych wartości skumulowanego udziału bankrutów, wyznaczonych dla grupy jednostek upadłych.

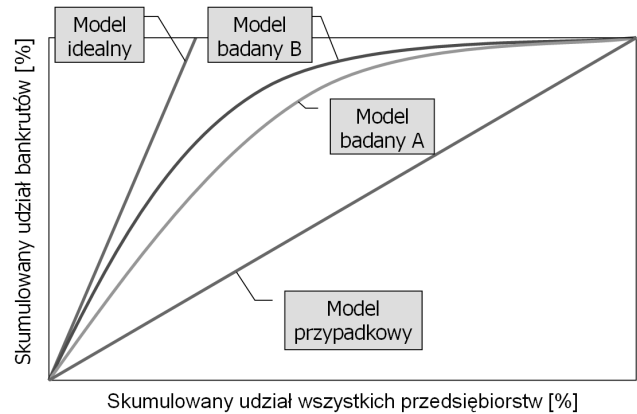
W układzie współrzędnych, w którym wykreślone zostają krzywe modeli badanych (modeli rzeczywistych), narysowane zostają ponadto 2 proste:

- prosta modelu idealnego (tj. takiego, który prawidłowo przewiduje sytuację przedsiębiorstw bez popelnienia żadnego błędu),
- prosta modelu przypadkowego (dającego najgorsze wyniki klasyfikacji).

Krzywe modeli rzeczywistych przebiegają pomiędzy prostymi modelu idealnego i przypadkowego (rys. 1).

Model jest tym dokładniejszy, im bliżej prostej modelu idealnego przebiega jego krzywa koncentracji CAP.

W poglądowym przykładzie pokazanym na rys. 1. krzywa modelu B znajduje się bliżej krzywej koncentracji modelu idealnego, w porównaniu z krzywą modelu A. Zatem, model B lepiej klasyfikuje przedsiębiorstwa do określonej grupy od modelu A.

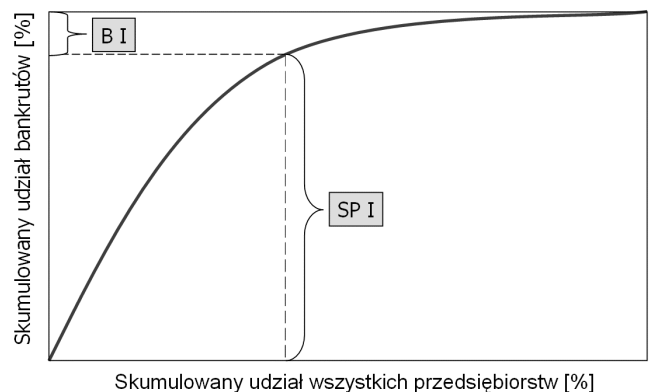


Rys.1. Krzywe koncentracji CAP

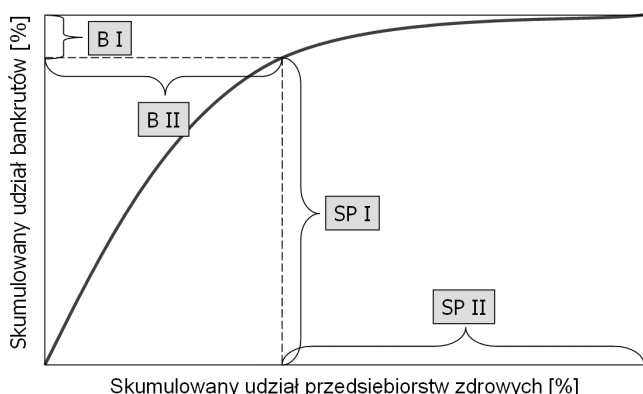
W przypadku krzywych koncentracji ROC, oś X jest opisana procentową skumulowaną wartością przedsiębiorstw niezagrożonych upadkiem. Pozostałe założenia są takie same, jak dla krzywych CAP.

Wyniki dostarczane przez krzywe koncentracji CAP i ROC można porównać następująco:

- krzywa CAP dostarcza odpowiedzi na pytanie, ile procent przedsiębiorstw trzeba wyłączyć z całej badanej grupy, aby wyłączyć określony procent bankrutów?
- z kolei krzywa ROC dostarcza odpowiedzi na pytanie, ile procent przedsiębiorstw trzeba wyłączyć z grupy jednostek niezagrożonych upadkiem, aby wyłączyć określony procent bankrutów?
- krzywa CAP pozwala odczytać wyłącznie sprawność i błąd I rodzaju (rys. 2).
- natomiast krzywa ROC pozwala odczytać sprawności i błędy I i II rodzaju (rys. 3).



Rys.2. Określanie błędu i sprawności I rodzaju na podstawie krzywej koncentracji CAP

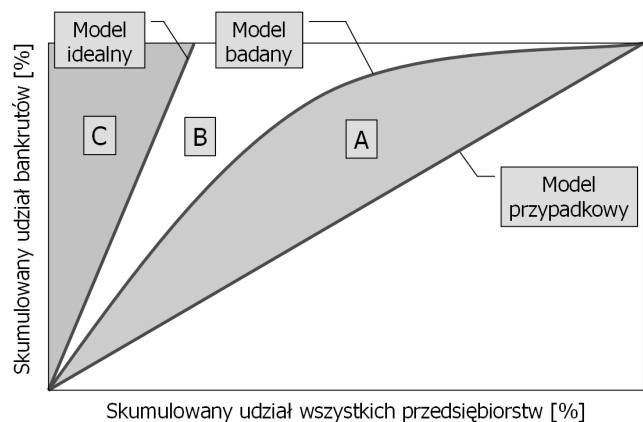


Rys.3. Określanie błędów i sprawności I i II rodzaju na podstawie krzywej koncentracji ROC

Krzywe koncentracji, a szczególnie ROC, mają tę przewagę nad macierzą klasyfikacji, że przedstawiają wyniki klasyfikacji dla wielu punktów granicznych.

2.4. Wskaźnik dokładności (wskaźnik Giniego)

Wskaźnik dokładności można wyznaczyć na podstawie obu krzywych, tj. CAP i ROC. Jego wartość należy do przedziału od 0 do 1, przy czym im jest ona wyższa, tym lepszy (dokładniejszy) jest badany model.



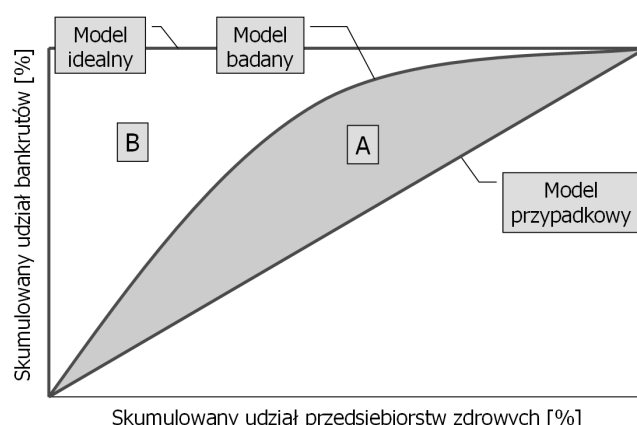
Rys.4. Wyznaczanie wskaźnika dokładności na podstawie krzywej koncentracji CAP

Korzystając z krzywej koncentracji CAP (rys. 4), wskaźnik dokładności obliczamy na podstawie następującej zależności:

$$(8) \quad WD = \frac{\text{Obszar A}}{\text{Obszar A} + \text{Obszar B} + \text{Obszar C}}$$

Korzystając z krzywej koncentracji ROC (rys. 5), wskaźnik dokładności obliczamy na podstawie następującej zależności:

$$(9) \quad WD = \frac{\text{Obszar A}}{\text{Obszar A} + \text{Obszar B}}$$



Rys.5. Wyznaczanie wskaźnika dokładności na podstawie krzywej koncentracji ROC

2.5. Zmodyfikowany wskaźnik dokładności

W przypadku, gdy porównywane są wyniki uzyskane z prób różniących się udziałem bankructw, wykorzystywany jest zmodyfikowany wskaźnik dokładności (ZWD):

$$(10) \quad ZWD = \frac{\text{Obszar A}}{\text{Obszar A} + \text{Obszar B}}$$

Pomiędzy wskaźnikiem dokładności, a zmodyfikowanym wskaźnikiem dokładności występuje zależność umożliwiającą ich wzajemne przeliczanie:

$$(11) \quad ZWD = \frac{WD}{1 - c}$$

gdzie: c – udział bankructw.

2.6. Ograniczenia krzywych koncentracji i wskaźnika dokładności

Krzywe koncentracji i wskaźnik dokładności nie pozwalają na porównanie różnic odległości wartości funkcji z-score sklasyfikowanych jednostek od wartości granicznej. Może również wystąpić pewien problem, związany z oszacowaniem dokładności modeli, ze względu na możliwość przecinania się krzywych koncentracji poszczególnych modeli. Problem ten można jednak rozwiązać kosztem bardziej złożonej metodyki postępowania.

3. Metody oceny kalibracji modeli

3.1. Wskaźnik Briera

Wskaźnik Briera (BS) obliczany jest według następującej formuły [1]:

$$(12) \quad BS = \frac{1}{n} \sum_{i=1}^n (\Theta_i - p_i)^2$$

gdzie:

n – liczba obserwacji w próbie,

i – poszczególne obserwacje,

Θ_i – zmienna, na podstawie której przypisuje się przed-

siębiorstwu 1, gdy jest ono w rzeczywistości bankrutem oraz 0, gdy nie jest zagrożone upadkiem,
 p_i – oszacowane prawdopodobieństwo upadku zarówno dla jednostek będących w rzeczywistości bankrutami, jak i niezagrażonych upadłością.

Im niższa jest wartość wskaźnika Briera, tym lepszy jest badany model. Wskaźnik Briera, w porównaniu z macierzą klasyfikacji, ilorazem szans oraz współczynnikiem dokładności, jest bardziej wrażliwy na poziom szacowanych wartości prawdopodobieństw (są one uwzględnione bezpośrednio).

3.2. Wskaźnik Briera dla prognozy przypadkowej

Modele można również porównywać wykorzystując wskaźnik Briera obliczony dla prognozy przypadkowej. Zakłada się wówczas, że prawdopodobieństwo określenia stanu upadku przedsiębiorstwa jest równe przeciętnemu udziałowi bankrutów w badanej próbie. Wskaźnik Briera dla prognozy przypadkowej (PBS) wyznaczany jest z poniższej zależności:

$$(13) \quad PBS = 1/n \left[\begin{array}{l} n_{DEF=1} \cdot (ADF-1)^2 \\ + n_{DEF=0} \cdot (ADF-0)^2 \end{array} \right]$$

gdzie:

n – liczba obserwacji w próbie,

$n_{DEF=1}$ – rzeczywista liczba bankrutów,

$n_{DEF=0}$ – rzeczywista liczba przedsiębiorstw niezagrażonych upadkiem,

ADF – przeciętne prawdopodobieństwo upadku przedsiębiorstwa wyznaczone dla danej próby badawczej.

Im niższa jest wartość wskaźnika Briera, w porównaniu z tą wartością dla prognozy przypadkowej, tym lepszy jest badany model.

3.3. Zmodyfikowany wskaźnik Briera

W przypadku, gdy próby różnią się udziałem bankrutów w łącznej liczbie przedsiębiorstw, stosowany jest zmodyfikowany wskaźnik Briera (ZBS). Wyznacza się go według następującej formuły:

$$(14) \quad ZBS = 1 - \frac{BS}{PBS}$$

gdzie:

ZBS – zmodyfikowany wskaźnik Briera,

BS – wskaźnik Briera badanego modelu,

PBS – wskaźnik Briera dla prognozy przypadkowej.

Im bardziej wartość ZBS jest zbliżona do jedności, tym lepszy jest badany model. Wskaźnik ZBS należy wyznaczać jedynie dla celów poglądowych, wówczas gdy porównywane są wyniki uzyskane na podstawie badań przeprowadzonych przy wykorzystaniu prób testowych o różnym udziale bankrutów w łącznej liczbie przedsiębiorstw.

3.4. Wskaźnik wiarygodności – L (model)

Oszacowane prawdopodobieństwo wystąpienia określonego stanu Y, przy danych wartościach zmiennych objaśniających X, można wyznaczyć następująco:

$$(15) \quad P(Y/X) = P(X)^Y \cdot [1 - P(X)]^{1-Y}$$

gdzie:

$P(X)$ – prawdopodobieństwo upadłości oszacowane za pomocą modelu, przy danych wartościach zmiennych objaśniających X,

Y – stan przyjmujący wartość 1, gdy przedsiębiorstwo jest w rzeczywistości bankrutem i 0, gdy jest niezagrażone upadkiem.

Wskaźnik wiarygodności obliczany jest według następującej zależności:

$$(16) \quad L(\text{model}) = \prod_{i=1}^n P(Y_i/X_i)$$

gdzie:

n – łączna liczba obserwacji,

i – poszczególne obserwacje,

$P(Y_i/X_i)$ – oszacowane prawdopodobieństwo wystąpienia określonego stanu, przy danych wartościach zmiennych objaśniających X, wyznaczone dla i -tej obserwacji.

Model charakteryzuje się tym lepszą kalibracją, im większa jest wartość wskaźnika wiarygodności.

Literatura

1. Brier G., Verification of forecasts expressed in terms of probability, Monthly Weather Review, 78:1,3, 1950, <http://docs.lib.noaa.gov/rescue/mwr/078/mwr-078-01-0001.pdf>.
2. James V., Herrity J. V., Keenan S., Sobehart J. R., Carty L. V., Falkenstein E. G., Measuring Private Firm Default Risk, Moody's KMV Technical Report, June 1999, <http://www.moody's.com/us/research/crm/45768.pdf>
3. Prusak B., Nowoczesne metody prognozowania zagrożenia finansowego przedsiębiorstw, Difin, Warszawa 2005.
4. Sobehart J. R., Keenan S. C., Stein R. M., Benchmarking Quantitative Default Risk Models: A Validation Methodology, Algo Research Quarterly, Vol. 4, No. 1/2, March/June 2001, <http://www.moody's.com/us/research/crm/53621.pdf>
5. Stein R. M., Ahmet E. Kocagil A. E., Bohn J. Akhavein J., Systematic And Idiosyncratic Risk In Middle-Market Default Prediction: A Study Of The Performance Of The RiskCalc™ And PFM™ Models, Moody's KMV Technical Report, February 2003, <http://www.moody's.com/us/research/crm/riskcalcpfm.pdf>
6. Stein R. M., Benchmarking Default Prediction Models: Pitfalls and Remedies in Model Validation, Moody's KMV Technical Report, June 2002, http://riskcalc.moody's.com/us/research/crm/Validation_Tech_Report_020305.pdf