

**Michał Trzęsiok**

Akademia Ekonomiczna w Katowicach

---

## WYODRĘBNIANIE REGUŁ KLASYFIKACYJNYCH Z MODELU DYSKRYMINACYJNEGO BUDOWANEGO METODĄ WEKTORÓW NOŚNYCH

---

**Streszczenie:** Metoda wektorów nośnych (SVM) jest jedną z najdokładniejszych metod dyskryminacji. Istotną barierą dla zwiększenia zakresu zastosowań metody SVM jest to, że otrzymywany model jest bardzo słabo interpretowalny. W celu wyodrębnienia reguł klasyfikacyjnych, według których metoda decyduje o wskazaniu przynależności do klasy, można wykorzystać modele drzew klasyfikacyjnych, których wyniki są łatwe w interpretacji. W artykule zaproponowano procedurę wyodrębniania reguł z modeli SVM i zilustrowano ją na danych rzeczywistych.

### 1. Wstęp

W ciągu ostatnich dwudziestu lat obserwowalny jest bardzo dynamiczny rozwój metod statystycznych wykorzystujących techniki automatycznego uczenia się (*machine learning systems*). Popularność oraz gwałtowny rozwój tej grupy metod wynika przede wszystkim z ich dominacji nad innymi metodami pod względem zdolności predykcyjnych odpowiadających im modeli. Jednym z narzędzi eksploracji danych, które w zagadnieniach dyskryminacji jest uważane za jedno z najdokładniejszych [Abe 2005; Trzęsiok 2006], jest zaproponowana przez Vapnika [1998] metoda wektorów nośnych (SVM – *Support Vector Machines*). Znaczną barierą dla zwiększenia zakresu zastosowań metody SVM jest to, że model otrzymywany w rezultacie jej zastosowania jest bardzo słabo interpretowalny. Choć modele SVM pozwalają na ogół na bardzo dokładną predykcję, to nie jest jednak możliwe zrozumiałe i przystępne wyjaśnienie w języku naturalnym reguł klasyfikacyjnych, według których model decyduje o wskazaniu przynależności do klasy dla pewnej nowej obserwacji. Z tego powodu metody eksploracji danych, w tym metodę SVM, określa się mianem czarnych skrzynek. W celu usunięcia tej poważnej wady wciąż pojawia się wiele propozycji modyfikacji lub nowych, niezależnych metod, które zastosowane po zbudowaniu modelu, mają za zadanie zwiększenie ich interpretowalności. Jednym z ważnych zagadnień związanych z osiąganiem tego celu jest analiza mocy dyskryminacyjnej zmiennych objaśniających oraz identyfikacja zmiennych nieistotnych [Guyon i in. 2006; Trzęsiok 2009]. Drugie ważne zagad-

nienie to wyodrębnienie reguł klasyfikacyjnych, któremu poświęcona jest niniejsza praca.

Zdolność do wyjaśnienia i zrozumiałego przedstawienia reguł klasyfikacyjnych jest bardzo ważnym elementem w każdym zagadnieniu, w którym zbudowany model dyskryminacyjny przed zastosowaniem musi zostać zweryfikowany przez grupę ekspercką, np. w przypadku modeli dyskryminacyjnych realizujących diagnostykę medyczną lub w przypadku modeli klasyfikujących klientów banku do grupy osób, którym można bezpiecznie udzielić kredytu lub nie.

Pierwsze próby i techniki wyodrębniania reguł z modeli SVM były inspirowane głównie metodami zaproponowanymi do tego samego celu dla sieci neuronowych [Diederich 2008]. Jedno z takich podejść wykorzystuje modele drzew klasyfikacyjnych do wygenerowania reguł klasyfikacyjnych dla danych, gdzie rzeczywiste wartości zmiennej definiującej klasę obiektu zastąpiono wskazaniem przynależności do klas otrzymanymi z modelu SVM. W ten sposób łatwe w interpretacji wyniki metody drzew klasyfikacyjnych pozwalają na zrealizowanie etapu profilowania klas otrzymanych metodą wektorów nośnych (wyodrębniania reguł klasyfikacyjnych). To połączenie dwóch różnych metod pozwala na wzbogacenie modelu SVM o pożądane własności, które mają drzewa klasyfikacyjne.

W części drugiej artykułu w syntetycznej formie przedstawiono metodę wektorów nośnych ze szczególnym uwzględnieniem wszystkich tych jej aspektów, które sprawiają, że końcowy model ma tak nieprzystępną dla interpretacji postać. Szczegółowy opis metody SVM można znaleźć m.in. w [Vapnik 1998; Abe 2005; Cristianini, Shawe-Taylor 2000; Trzysiak 2004]. W trzecim punkcie artykułu omówiona została metoda wyodrębniania reguł wykorzystująca drzewa klasyfikacyjne, w czwartej zaś części przedstawiono przykład ilustrujący zaprezentowaną procedurę na zbiorze danych rzeczywistych.

## 2. Syntetyczny opis metody wektorów nośnych

Rozważmy przypadek zadania dyskryminacji, w którym obserwacje należą do jednej z dwóch klas. W przypadku większej liczby klas można wyznaczyć wiele funkcji dyskryminujących klasy parami i skonstruować funkcję dyskryminującą, korzystając z reguły majoryzacji (głosowania modeli cząstkowych). Niech zatem dany będzie zbiór uczący  $D = \{(x^1, y^1), \dots, (x^N, y^N)\}$ , gdzie  $x^i \in \mathbf{R}^d$  to wektor zmiennych objaśniających, a  $y^i \in \{-1, 1\}$  – wartości zmiennej opisującej klasę obiektu. Zadanie dyskryminacji polega na skonstruowaniu hiperpowierzchni rozdzielającej obiekty należące do dwóch różnych klas na podstawie informacji zawartej w zbiorze uczącym  $D$ . Wyznaczonej hiperpowierzchni odpowiada funkcja klasyfikująca dowolne obiekty z przestrzeni  $\mathbf{R}^d$ .

W celu zredukowania złożoności zadania wyznaczania hiperpowierzchni rozdzielającej klasy metoda wektorów nośnych wykorzystuje funkcje liniowe (hiper-

plaszczyny). Żeby jednak umożliwić również nieliniową klasyfikację, hiperplaszczynami rozdzielane są nie obserwacje ze zbioru pierwotnego, lecz ich obrazy w przestrzeni o znacznie większym (czasem nieskończonym) wymiarze. Mianowicie obserwacje ze zbioru uczącego są transformowane za pomocą nieliniowej transformacji  $\varphi$  w nową przestrzeń cech  $\mathbf{Z}$ . Zadanie dyskryminacji polega więc na wyznaczeniu hiperplaszczyny

$$\boldsymbol{\beta} \cdot \varphi(\mathbf{x}) + \beta_0 = 0, \quad (1)$$

rozdzielającej klasy zmodyfikowanego zbioru uczącego  $\{(\varphi(\mathbf{x}^1), y^1), \dots, (\varphi(\mathbf{x}^N), y^N)\}$ ,

gdzie  $\varphi(\mathbf{x}^i) \in \mathbf{Z}$  oraz  $y^i \in \{-1, 1\}$  dla  $i = 1, \dots, N$ . Wobec nieliniowości transformacji  $\varphi$  liniowej funkcji rozdzielającej klasy w  $\mathbf{Z}$  odpowiada nieliniowa funkcja dyskryminująca w przestrzeni  $\mathbf{R}^d$ . Jeśli rozważane klasy są liniowo separowane w przestrzeni  $\mathbf{Z}$ , to zadanie wyznaczenia hiperplaszczyny rozdzielającej (1) nie ma jednoznacznego rozwiązania – istnieje nieskończenie wiele takich hiperplaszczyn. W celu wyznaczenia hiperplaszczyny dającej największe prawdopodobieństwo poprawnego klasyfikowania obserwacji spoza zbioru uczącego poszukiwana jest *hiperplaszczyna optymalna*, tzn. taka, dla której odległość od najbliższych obserwacji zarówno jednej, jak i drugiej klasy jest największa.

Postawione zadanie wyznaczenia hiperplaszczyny optymalnej można zapisać w postaci zadania optymalizacji wypukłej z kwadratową funkcją celu i liniowymi ograniczeniami reprezentującymi warunek rozdzielania klas [Cristianini, Shawe-Taylor 2000; Abe 2005]. Owo zadanie optymalizacyjne można rozwiązać metodą mnożników Lagrange’a, rozwiązanie zaś ma postać:

$$f(\mathbf{x}) = \text{sign} \left[ \sum_{i=1}^N \alpha_i y^i \varphi(\mathbf{x}^i) \cdot \varphi(\mathbf{x}) + \hat{\beta}_0 \right] = \text{sign} \left[ \sum_{i \in I_{SV}} \alpha_i y^i \varphi(\mathbf{x}^i) \cdot \varphi(\mathbf{x}) + \hat{\beta}_0 \right]. \quad (2)$$

W zadaniu optymalizacyjnym występuje bardzo wiele mnożników Lagrange’a  $\alpha_i$ . Jest ich tyle, ile liniowych ograniczeń, czyli tyle, ile obserwacji w zbiorze uczącym. Z warunków optymalności Karusha–Kuhna–Tuckera wynika, że wiele współczynników Lagrange’a w rozwiązaniu zadania optymalizacyjnego jest równych zero. Postać funkcji dyskryminującej (2) zależy więc wyłącznie od tych wektorów  $\mathbf{x}^i$  (obserwacji ze zbioru uczącego), którym odpowiadają niezerowe współczynniki Lagrange’a (zbiór indeksów tych wektorów został oznaczony we wzorze (2) symbolem  $I_{SV}$ ). Obserwacje te nazywamy *wektorami nośnymi*. W celu uniknięcia problemów obliczeniowych oraz merytorycznych związanych ze sposobem definiowania nieliniowej transformacji  $\varphi$  metoda wektorów nośnych wykorzystuje

funkcje jądrowe, o których wiadomo, że definiują iloczyn skalarny w pewnej przestrzeni cech. Tym samym postać funkcji transformującej  $\varphi$  nie musi być znana, albowiem zarówno w trakcie rozwiązywania wspomnianego zadania optymalizacyjnego, jak i do zapisu jego rozwiązania (2) wystarczy mieć zdefiniowaną postać iloczynu skalarnego  $K(\mathbf{u}, \mathbf{v}) = \varphi(\mathbf{u}) \cdot \varphi(\mathbf{v})$  w przestrzeni  $\mathbf{Z}$ . W metodzie wektorów nośnych najczęściej wykorzystuje się funkcje jądrowe Gaussa, wielomianowe, sigmoidalne lub liniowe (zob. [Trzęsiok 2004]).

Zakładając, że w zbiorze uczącym mogą się znajdować obserwacje błędnie sklasyfikowane lub wartości zmiennych zaburzone szumem, Vapnik wprowadził do algorytmu metody pomocnicze, zmienne pozwalające na wygenerowanie rozwiązania, w którym niektóre obserwacje ze zbioru uczącego znajdują się po „niewłaściwej stronie” wyznaczonej hiperpłaszczyzny rozdzielającej klasy. Tym samym metoda SVM została wzbogacona o mechanizm eliminujący efekt nadmiernego dopasowania modelu dyskryminacyjnego do danych ze zbioru uczącego. Do zrealizowania metody SVM konieczne jest podanie kilku parametrów: jeden z nich określa kompromis między dokładnością dopasowania modelu do danych ze zbioru uczącego a zdolnością tej funkcji do poprawnego klasyfikowania nowych obserwacji. Pozostałe parametry wiążą się z wyborem typu funkcji jądrowej. Zagadnienie wyboru tychże parametrów zostało szczegółowo przedstawione w pracy [Trzęsiok 2004].

Wobec powyższych uwag funkcję dyskryminującą (2), będącą rozwiązaniem rozważanego zadania dyskryminacji, można ostatecznie zapisać w postaci:

$$f(\mathbf{x}) = \text{sign} \left[ \sum_{i \in I_{SV}} \alpha_i y^i K(\mathbf{x}^i, \mathbf{x}) + \hat{\beta}_0 \right]. \quad (3)$$

Zauważmy, że bardzo trudno z postaci modelu dyskryminacyjnego (3) jest wyodrębnić reguły przynależności do danej klasy, ponieważ identyfikowanie przynależności do klas odbywa się w przestrzeni o znacznie większym wymiarze niż pierwotna przestrzeń danych. Choć sam zapis modelu (3) ma dosyć zwartą postać, to jednak wykorzystując funkcje jądrowe  $K$ , definiujące iloczyn skalarny w pewnej przestrzeni, tracimy możliwość bezpośredniego kontrolowania, jak tworzone są nowe zmienne opisujące obiekty. Bardzo często nie można *explicite* podać postaci analitycznej funkcji transformującej  $\varphi$ . Ponadto liczba wektorów nośnych może być znaczna (liczba składników w (3)), co dodatkowo zwiększa złożoność postaci modelu i utrudnia odczytanie reguł klasyfikujących.

### 3. Procedura wyodrębniania reguł z modelu SVM

Istnieje kilka technik wyodrębniania reguł z modeli SVM na podstawie analizy wektorów nośnych – obserwacji identyfikowanych w modelu. Tak np. Núñez i in. w pracy [Diederich 2008] wskazują sposób wyodrębniania reguł z modeli SVM

przez zastosowanie metody  $k$ -średnich – do pogrupowania wektorów nośnych, by następnie dla każdego skupienia osobno wyznaczyć *prototyp*, tj. środek ciężkości skupienia. W kolejnym kroku procedury w obrębie danego skupienia wyznacza się wektor nośny najbardziej oddalony od prototypu i na podstawie tych dwóch obserwacji buduje się elipsoidę, która jest transformowana w zbiór reguł typu „jeżeli ..., to ...”. Powyższa metoda wyodrębniania reguł, jak też grupa metod uwzględniająca budowanie hiperprostopadłościanów wokół prototypów zamiast elipsoid mają poważną wadę, ponieważ są niezrozumiałe dla ekspertów weryfikujących model, gdyż bazują w swoim opisie na wartościach numerycznych wziętych z wektorów nośnych, co znacznie zmniejsza przejrzystość reguł.

Bardzo prostą techniką wyodrębniania reguł jest próba wykorzystania innych metod dyskryminacji, dających modele łatwo interpretowalne, do „nałożenia” na wynik działania metody SVM w taki sposób, żeby naśladowały mechanizm klasyfikowania zawarty w metodzie SVM (jak najlepiej imitowały działanie „czarnej skrzynki”). Kilka takich propozycji wykorzystujących m.in. metody C4.5, RIPPER, metodę samoorganizującej się sieci zawarł Martens i in. w pracy [Diederich 2008]. Również w opracowaniu [Diederich 2008] Pang i Kasabov przedstawili propozycję techniki wyodrębniania reguł z zagregowanych modeli SVM połączonych z drzewami klasyfikacyjnymi. W niniejszej pracy wykorzystane zostanie prostsze podejście polegające na zbudowaniu pojedynczego drzewa klasyfikacyjnego na zbiorze danych, w którym przynależność do klas została uzyskana z pojedynczego modelu SVM. W ten sposób zadaniem drzewa klasyfikacyjnego będzie imitowanie mechanizmu dyskryminacji zawartego w metodzie wektorów nośnych. Jeżeli tylko model drzewa będzie w stanie dobrze naśladować klasyfikację metody SVM (wysoka zgodność klasyfikacji metody SVM i drzewa), to uzyskane połączenie modeli będzie miało pożądane własności – wszystkie zalety stosowania metody SVM: dużą dokładność predykcji, możliwość nieliniowego rozdzielania klas, odporność na szum i obserwacje oddalone [Trzęsiok 2006; 2008] oraz dwie bardzo ważne własności modelu drzewa klasyfikacyjnego: łatwą interpretowalność reguł klasyfikacyjnych zapisanych w języku naturalnym (reguły koniunkcyjne) oraz możliwość graficznej prezentacji modelu w postaci grafu (drzewa) [Gatnar 2001].

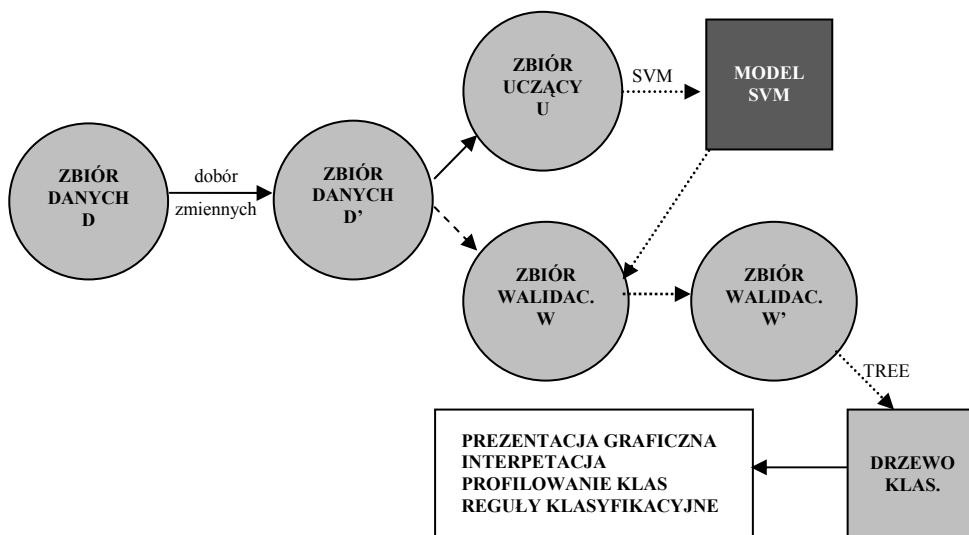
Powód zastosowania drzew klasyfikacyjnych wraz z metodą wektorów nośnych, a nie samej metody drzew klasyfikacyjnych, jest taki, że jak wskazują badania symulacyjne – pojedyncze drzewo, choć ma wiele pożądanych własności, to jednak na ogół znacznie ustępuje pod względem zdolności predykcyjnych (dokładności klasyfikacji) metodzie SVM [Trzęsiok 2006]. Metoda wektorów nośnych jest więc niejako wykorzystana do oczyszczenia analizowanego zbioru danych z szumu, z negatywnego wpływu występowania obserwacji nietypowych, do rozpoznania nieliniowych postaci hiperpowierzchni rozdzielających klasy w sposób umożliwiający możliwie jak najdokładniejsze klasyfikowanie obserwacji spoza zbioru uczącego. Tak skonstruowane narzędzie o dobrych własnościach formalno-statystycznych zostanie następnie wykorzystane do sklasyfikowania nowych ob-

serwacji ze zbioru walidacyjnego. Po zastąpieniu rzeczywistej przynależności do klas obiektów ze zbioru walidacyjnego wskazaniem z modelu SVM na tak zmodyfikowanym zbiorze walidacyjnym zostanie zbudowane pojedyncze drzewo, z którego następnie odczytany zostanie zbiór reguł klasyfikujących. Dodatkowo, aby w opisie reguł nie uwzględniać informacji mało istotnych, zwiększających jedynie złożoność modelu, przed zbudowaniem modelu SVM zastosowana zostanie proce-

**Tabela 1.** Procedura wyodrębniania reguł z modeli SVM

|                |  |
|----------------|--|
| <b>Krok 1.</b> | <p><b>Usunięcie zmiennych redundantnych:</b></p> <p>a) Zbuduj model SVM na zbiorze uczącym <math>D</math>, wykorzystując pełen zestaw zmiennych i dobierając optymalnie wartości parametrów metody. Utwórz pomocniczy zbiór uczący <math>S</math> będący kopią zbioru <math>D</math>.</p> <p>b) Wygeneruj wiele zmodyfikowanych zbiorów uczących na bazie <math>S</math>, wyłączając za każdym razem tymczasowo jedną zmienną z pierwotnego zbioru <math>S</math>. Następnie na tych zbiorach zbuduj modele SVM z parametrami identycznymi jak w punkcie a).</p> <p>c) Oblicz metodą sprawdzania krzyżowego błąd klasyfikacji <math>CVerr</math> każdego z modeli z punktu b).</p> <p>d) Zidentyfikuj model z wyłączoną zmienną, dla którego wartość błędu <math>CVerr</math> jest najmniejsza. Usuń ostatecznie ze zbioru <math>S</math> zidentyfikowaną zmienną.</p> <p>e) Powtarzaj kroki b)-d), dopóki w <math>S</math> pozostanie jedna zmienna.</p> <p>f) Z sekwencji modeli z zestawem zmiennych o liczebności od <math>d</math> do 1 wybierz model z najmniejszą liczbą zmiennych, którego błąd jest nie większy niż <math>\min(CVerr) + SE</math>, czyli minimalny błąd klasyfikacji powiększony o błąd standardowy pomiaru. Zmienne nieuwzględnione w tym modelu uznaj za redundantne i usuń je ostatecznie ze zbioru uczącego <math>D</math>, tworząc zbiór <math>D'</math> (zob. [Trzęsiok 2009]).</p> |
| <b>Krok 2.</b> | Podziel zbiór $D'$ na część uczącą $U$ (50%) i walidacyjną $W$ (50%).  |
| <b>Krok 3.</b> | <p><b>Uczenie modelu SVM relacji w analizowanym zbiorze:</b></p> <p>Zbuduj model SVM na części uczącej <math>U</math>, dobierając optymalne wartości parametrów metodą sprawdzania krzyżowego.</p>   |
| <b>Krok 4.</b> | <p><b>Wykorzystanie wiedzy zawartej (ukrytej) w modelu SVM do klasyfikowania nowych obiektów:</b></p> <p>poddaj klasyfikacji obserwacje ze zbioru walidacyjnego <math>W</math>, wykorzystując model zbudowany w kroku 3. Podmień rzeczywistą przynależność do klas obserwacji ze zbioru walidacyjnego <math>W</math> na wskazania przynależności do klas otrzymane z modelu SVM, tworząc nowy zbiór <math>W'</math>.</p>   |
| <b>Krok 5.</b> | <p><b>Budowanie drzewa klasyfikacyjnego naśladującego działanie modelu SVM (uczenie drzewa tego, czego nauczył się model SVM):</b></p> <p>zbuduj drzewo klasyfikacyjne na zbiorze <math>W'</math>. Sprawdź zgodność klasyfikacji tak zbudowanego drzewa ze wskazaniami modelu SVM (na ile drzewo klasyfikacyjne dobrze imituje ukryty mechanizm dyskryminacji zawarty w SVM).</p>  |
| <b>Krok 6.</b> | <p><b>Odczytanie reguł klasyfikacyjnych dla modelu SVM ze zbudowanego drzewa klasyfikacyjnego naśladującego działanie modelu SVM:</b></p> <p>jeśli zgodność klasyfikacji drzewa ze wskazaniami modelu SVM jest względnie wysoka (ocena subiektywna), wykorzystaj reguły klasyfikacyjne otrzymane z drzewa klasyfikacyjnego oraz jego reprezentację graficzną jako opis w etapie profilowania klas modelu SVM</p>   |

Źródło: opracowanie własne.



Rys. 1. Schemat procedury wyodrębniania reguł z modelu SVM

Źródło: opracowanie własne.

dura identyfikacji zmiennych istotnych w modelu SVM opisana w [Trześciok 2009]. Zmienne redundantne są usuwane ze zbioru uczącego.

Całą procedurę wyodrębniania reguł z uwzględnieniem eliminacji zmiennych nieistotnych przedstawiono w tab. 1 oraz schematycznie na rys. 1.

#### 4. Przykład ilustrujący procedurę wyodrębniania reguł z modelu SVM

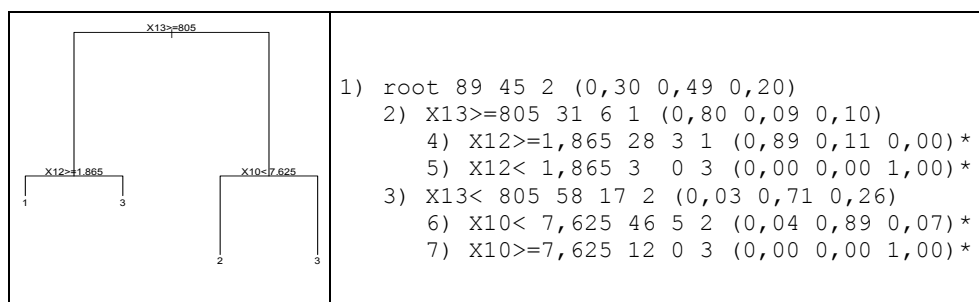
Przedstawiona procedura wyodrębniania reguł zilustrowana zostanie na zbiorze danych rzeczywistych *Wine* wykorzystywanym do badania własności metod wielowymiarowej analizy statystycznej. Zbiór ten dostępny jest w repozytorium Uniwersytetu Kalifornijskiego<sup>1</sup>, zawiera 178 obserwacji charakteryzowanych przez 13 zmiennych objaśniających, należących do jednej z 3 klas. Wszystkie obliczenia zostały zrealizowane z wykorzystaniem programu statystycznego **R** i jego dodatkowych pakietów: `e1071`, zawierającego funkcję realizującą metodę SVM, oraz `rpart` – dla metody drzew klasyfikacyjnych.

W pierwszym kroku procedury opisanej w tab. 1 zidentyfikowano cztery zmienne redundantne o numerach: 1, 5, 7 i 11, które zostały usunięte z rozważanego zbioru danych i nie uczestniczyły w dalszej analizie. Tak zmodyfikowany zbiór podzielono losowo na dwie części. Część ucząca została wykorzystana do zbudowania

<sup>1</sup> Udostępniany pod adresem: <ftp://ftp.ics.uci.edu/pub/machine-learning-databases>.

wania modelu SVM, który następnie wykorzystano do sklasyfikowania obserwacji ze zbioru walidacyjnego. Zbudowano model drzewa klasyfikacyjnego na zbiorze walidacyjnym ze wskazaniem przynależności do klas wynikającymi z modelu SVM. Zgodność klasyfikacji drzewa oraz wskazań modelu SVM na zbiorze walidacyjnym jest równa 91%. Można więc przyjąć, że drzewo klasyfikacyjne w bardzo dużym stopniu naśladuje sposób klasyfikacji ukryty w modelu SVM.

Na rysunku 2 przedstawiono graficzną reprezentację drzewa naśladującego wskazania modelu SVM oraz postać modelu w postaci tekstowej generowaną w programie **R**.



**Rys. 2.** Postać graficzna oraz tekstowa modelu drzewa klasyfikacyjnego naśladującego reguły klasyfikacji zawarte w modelu SVM

Źródło: opracowanie własne.

Na rysunku 2 widać, że reguły klasyfikacji modelu SVM zostały sformułowane z wykorzystaniem drzewa klasyfikacyjnego za pomocą zmiennych  $X_{13}$ ,  $X_{12}$  oraz  $X_{10}$  i są postaci:

Jeśli ( $X_{13} \geq 805$  i  $X_{12} \geq 1,865$ ), to dany obiekt przynależy do klasy 1.

Jeśli ( $X_{13} < 805$  i  $X_{10} < 7,625$ ), to dany obiekt przynależy do klasy 2.

Jeśli  $\{(X_{13} \geq 805 \text{ i } X_{12} < 1,865) \text{ lub } (X_{13} < 805 \text{ i } X_{10} \geq 7,625)\}$ , to dany obiekt przynależy do klasy 3.

## 5. Podsumowanie

W artykule zaproponowano i zilustrowano na przykładzie danych rzeczywistych metodę wyodrębniania reguł z modeli SVM. Metoda ta łączy zalety modeli SVM (nieliniowość, odporność na szum, dobre właściwości predykcyjne) z zaletami drzew klasyfikacyjnych (reguły w prostej do interpretowania postaci, możliwość prezentacji graficznej modelu). Przedstawiona procedura pozwala na ekstrakcję wiedzy o badanym zjawisku z modeli SVM, które mają charakter czarnej skrzynki i trudno poddają się interpretacji bez użycia dodatkowych narzędzi. Wadą zaprezentowanego podejścia jest to, że do opisu działania mechanizmu klasyfikacji



w modelach SVM wykorzystuje niezależne narzędzie – metodę drzew klasyfikacyjnych, której mechanizm rozpoznawania relacji w zbiorze obserwacji jest całkowicie odmienny. Nawet w przypadku wysokiej zgodności wskazań modelu SVM i drzewa można twierdzić jedynie, że uzyskane z drzewa klasyfikacyjnego reguły klasyfikacji dobrze naśladują mechanizmy zawarte w modelu SVM, lecz ich nie reprezentują.

## Literatura

- Abe S., *Support Vector Machines for Pattern Classification*, Springer, 2005.
- Cristianini N., Shawe-Taylor J., *An Introduction To Support Vector Machines (and Other Kernel-Based Learning Methods)*, Cambridge University Press, Cambridge 2000.
- Diederich J. (red.), *Rule Extraction from Support Vector Machines*, Springer, 2008.
- Gatnar E., *Nieparametryczna metoda dyskryminacji i regresji*, PWN, Warszawa 2001.
- Guyon I., Gunn S., Nikravesh M., Zadeh L. (red.), *Feature Extraction, Foundations and Applications*, Springer 2006.
- Trzęsiok M., *Analiza wybranych własności metody dyskryminacji wykorzystującej wektory nośne*, [w:] *Postępy ekonometrii*, A.S. Barczak (red.), AE, Katowice 2004.
- Trzęsiok M., *Metoda wektorów nośnych na tle innych metod wielowymiarowej analizy danych*, [w:] *Taksonomia 13*, K. Jajuga, M. Walesiak (red.), AE, Wrocław 2006.
- Trzęsiok M., *Empiryczna ocena wrażliwości metody wektorów nośnych na występowanie obiektów błędnie sklasyfikowanych w zbiorze uczącym*, [w:] *Studia Ekonomiczne nr 50*, J. Mika (red.), AE, Katowice 2008.
- Trzęsiok M., *Problem doboru zmiennych do modelu dyskryminacyjnego budowanego metodą wektorów nośnych*, [w:] *Taksonomia 16*, K. Jajuga, M. Walesiak (red.), Wydawnictwo Uniwersytetu Ekonomicznego we Wrocławiu, Wrocław 2009.
- Vapnik V., *Statistical Learning Theory*, John Wiley&Sons, N.Y. 1998.

## RULES EXTRACTION FROM SUPPORT VECTOR MACHINES USING CLASSIFICATION TREES

**Summary:** Support Vector Machines (SVM) usually provide very accurate models. However, their main weakness is the lack of interpretation. Models generated by SVMs are regarded as very incomprehensible. In order to extract knowledge from SVMs, we use a simple technique – we combine SVMs with classification trees, which produces very comprehensible classification rules.