

Michał Trzęsiok

Uniwersytet Ekonomiczny w Katowicach
e-mail: michal.trzesiok@ue.katowice.pl

O WZBOGACANIU METOD KLASYFIKACJI W ZDOLNOŚĆ DO WYRAŻANIA WĄTPLIWOŚCI W PRZYPADKACH TRUDNYCH DO ROZSTRZYGNIECIA

PROVIDING CLASSIFICATION METHODS WITH THE ABILITY TO HESITATE IN CASES DIFFICULT TO SOLVE

DOI: 10.15611/pn.2017.469.22

JEL Classification: C150, C180, C300, C380

Streszczenie: Podejmując decyzje, czasem wahamy się, ponieważ czujemy, że w danej sytuacji wskazanie nie jest jednoznaczne. Z analogiczną sytuacją mamy do czynienia w statystycznych metodach eksploracyjnych. Metoda wykorzystuje informacje zawarte w zbiorze uczącym, model zostaje zbudowany, a następnie wykorzystany do predykcji na nowych obserwacjach. Wydaje się jednak rozsądne oczekiwać, żeby metoda dawała przynajmniej ostrzeżenie, kiedy wynik predykcji jest niestabilny, w tym sensie, że jest wrażliwy na małe zmiany w wartościach zmiennych objaśniających. Głównym celem artykułu jest przedstawienie procedury wzbogacającej metody klasyfikacji w zdolność do wskazywania, czy klasyfikacja danego obiektu jest trudna do rozstrzygnięcia. Przedstawiona procedura wykorzystuje analizę wrażliwości i jest uniwersalna, tzn. może być stosowana z różnymi metodami klasyfikacji. Dodatkowym walorem procedury jest zaproponowana w niej metoda wizualizacji wyników przeprowadzonej analizy wrażliwości.

Słowa kluczowe: metody dyskryminacji, analiza wrażliwości, *SVM*.

Summary: When making decision sometimes we hesitate, because we feel it can go both directions. The situation is similar in machine-learning tasks. We can use very sophisticated classification methods to support our decision-making process. The machine is learned, the model is built, but then it seems reasonable to expect the machine to give us at least a warning when the prediction is unstable (which means that it is sensitive to small changes in explanatory variables' values). The main goal of the article is to present a procedure for providing the machine with the ability to show hesitation, when it is justified. The proposed procedure is based on sensitivity analysis. We illustrate the procedure on a real-world data set using the Support Vector Machines, but the procedure is universal and it can be also used with other classification methods. The added value of the paper is also the proposed type of plot for visualizing the outcome of the sensitivity analysis.

Keywords: classification methods, sensitivity analysis, *SVMs*.

1. Wstęp

Statystyczne metody automatycznego uczenia się są z założenia konstruowane tak, by imitować proces odnajdowania reguł i relacji między analizowanymi cechami na podstawie pewnego zbioru przykładów. Proces ten jest odwzorowaniem sposobu uczenia się u ludzi. Reguły, których nauczyliśmy się i którymi posługujemy się, podejmując decyzje, nie zawsze dają jednoznaczne wskazanie, a nawet jeśli dają, to mamy czasem świadomość, że niewielka zmiana w uwarunkowaniach spowodowałaby podjęcie innej decyzji. W takich sytuacjach podjęcie decyzji jest trudnym zadaniem. Jest tak niezależnie od złożoności problemu, czy liczebności zbioru wariantów decyzyjnych. Nawet w przypadku decyzji z dwoma wariantami „tak” lub „nie”, czasem stajemy wobec sytuacji, w której odpowiadamy „trudno orzec”. Dodatkowo nie zawsze jest możliwe rozwiązanie tego problemu przez pozyskanie dodatkowych informacji. Nadmierne zwiększanie ilości informacji przy podejmowaniu decyzji może nawet prowadzić do zwiększenia problemów z ich podjęciem. Można przyjąć, że istnieją przypadki, w których wątpliwości co do podjęcia decyzji są ich immanentną cechą. Skoro więc wątpliwości są nieodłącznym i nieredukowalnym elementem procesów podejmowania decyzji, to wymaganie od statystycznych metod automatycznego uczenia się, by dawały zawsze jednoznaczne rozstrzygnięcia, nie wydaje się czymś rozsądnym. Nawet w bardzo wyrafinowanych metodach klasyfikacji danych wydaje się racjonalne oczekiwanie, by metoda przynajmniej dawała ostrzeżenie w przypadku, gdy klasyfikacja dla danej obserwacji jest niestabilna. Wyjaśnienia wymaga pojęcie *niestabilnej klasyfikacji*. W tym opracowaniu przyjęto, że niestabilna klasyfikacja jest definiowana jako takie wskazanie przynależności do klasy, które jest wrażliwe na małe odchylenia w wartościach zmiennych objaśniających. Z uwagi na to, że nie jest możliwe zdefiniowanie pojęcia *małych odchyleń* z wykorzystaniem wielkości absolutnych, bez odnoszenia się do konkretnego przypadku rozpatrywanej zmiennej, w dalszej części artykułu i w zaproponowanej procedurze wykorzystano względne odchylenia od oryginalnych wartości zmiennych objaśniających. Naturalną konsekwencją takiego podejścia jest to, że procedura może być wykorzystana wyłącznie dla zmiennych objaśniających, mierzonych na najmocniejszej skali – skali ilorazowej.

Głównym celem artykułu jest zaproponowanie procedury wzbogacającej metody klasyfikacji danych w zdolność do wskazywania, czy klasyfikacja danego obiektu wzbudza wątpliwości (tzn. jest trudna do rozstrzygnięcia) w przypadkach, w których jest to uzasadnione.

2. Procedura wzbogacania metod klasyfikacji w zdolność do wyrażania wątpliwości

Znane są metody klasyfikacji, które pozwalają na modelowanie niepewności, wykorzystując pojęcie zbiorów rozmytych lub zbiorów przybliżonych [Bellman i in. 1966; Pawlak 1982; Hong, Choi 2000]. Dominującą rolę w grupie wielowymiarowych me-

tod klasyfikacji pełnią jednak metody eksploracyjne, takie jak metoda *Random Forest*, *SVM*, sieci neuronowe czy *k*-najbliższych sąsiadów [Trzęsiok 2006]. Procedura zaproponowana w dalszej części opracowania ma z założenia jedynie uzupełniać działanie znanych metod klasyfikacji o dodatkowy element związany z uwzględnieniem niepewności decyzji klasyfikacyjnej, zachowując przy tym wszystkie właściwości oryginalnych metod. Nie jest to więc propozycja nowej metody klasyfikacji.

Konstrukcję procedury zrealizowano przy następujących ogólnych założeniach:

a) procedura ma realizować główny cel, czyli ma wzbogacać działanie metody klasyfikacyjnej w ten sposób, że ma dawać ostrzeżenie w przypadku, gdy wskazanie do klasy jest niestabilne,

b) procedura ma być prosta, o łatwo interpretowalnych wynikach oraz „lekka” obliczeniowo,

c) preferowane jest rozwiązanie umożliwiające wizualizację wyników działania procedury.

Przy takich założeniach skonstruowano procedurę w oparciu o analizę wrażliwości, przy czym spośród różnych metod analizy wrażliwości dla danych wielowymiarowych wykorzystano podejście „każdy wymiar z osobna” (tj. według zasady *ceteris paribus*, czyli badana jest wrażliwość wyników klasyfikacji przy wprowadzanych małych odchyleniach dla wybranej zmiennej, przy ustalonych, oryginalnych wartościach wszystkich pozostałych zmiennych objaśniających). Całą procedurę wzbogacania metod klasyfikacji w modelowanie niepewności przedstawiono w tabeli 1.

Tabela 1. Procedura wzbogacania metod klasyfikacji w modelowanie niepewności

Krok 1.	W uzasadnionych przypadkach dokonaj selekcji zmiennych lub przeprowadź inną procedurę wstępnego przetwarzania danych
Krok 2.	Zbuduj model klasyfikacyjny na danym zbiorze uczącym
Krok 3.	Zidentyfikuj zmienne objaśniające mierzone na skalach ilorazowych
Krok 4.	Ustal wartość progową – co uznajesz za małe odchylenie od oryginalnych wartości empirycznych (np. nie więcej niż 5%)
Krok 5.	Dla nowej obserwacji poddawanej klasyfikacji (ze zbioru rozpoznawanego) wyznacz wskazanie przynależności do klasy
Krok 6.	Utwórz dla każdej obserwacji ze zbioru rozpoznawanego zbiór nowych punktów w ten sposób, że mając pojedynczą obserwację, dla każdej zmiennej ilorazowej zwiększ (oraz zmniejsz) jej wartość w ustalonym zakresie (np. o 1%,..., 5% oryginalnej wartości empirycznej; w ten sposób, w przypadku gdy w zbiorze rozpoznawanych występują trzy zmienne objaśniające ze skali ilorazowej, z każdej jednej obserwacji ze zbioru rozpoznawanego powstaje 30 nowych obserwacji niewiele różniących się od obserwacji oryginalnej – po 10 obserwacji odchylających się o $\{-5\%, \dots, -1\%, +1\%, \dots, +5\%$ od wartości oryginalnej dla każdej z trzech zmiennych ilorazowych)
Krok 7.	Użyj modelu zbudowanego w kroku 2. do klasyfikacji nowych, zmodyfikowanych obserwacji skonstruowanych w kroku 6.

Tabela 1, cd.

Krok 8.	Jeśli wskazanie przynależności oryginalnej obserwacji różni się od wskazania dla którejkolwiek zmodyfikowanej, oznacz ją etykietą „wątpliwość” (obserwacja z dużą niepewnością predykcji); przedstaw wizualizację działania procedury na wykresach (dla każdej obserwacji z etykietą „wątpliwość” osobny wykres), na których różne figury geometryczne będą reprezentowały różne wskazania przynależności do klas, a kolejne poziomy na osi rzędnych będą reprezentować względne odchylenia (wyrażone w procentach) od oryginalnej wartości ilorazowej zmiennej objaśniającej
---------	---

Źródło: opracowanie własne.

Skonstruowanie procedury przy tak przyjętych założeniach jest problemem z obszaru ekstrakcji wiedzy z modelu i skutkuje zwiększeniem interpretowalności wyników działania metody klasyfikacji, co jest cennym elementem, zważywszy, że metody automatycznego uczenia się na ogół budują modele działające na zasadzie czarnej skrzynki i są słabo interpretowalne.

3. Przykład ilustrujący przedstawioną procedurę wzbogacania metod klasyfikacji w modelowanie niepewności

Przedstawiona procedura wzbogacania metod klasyfikacji w zdolność do wyrażania wątpliwości w przypadkach trudnych do rozstrzygnięcia zilustrowana zostanie na zbiorze danych rzeczywistych `german credit`, wykorzystywanym do badania własności metod wielowymiarowej analizy statystycznej. Zbiór ten – udostępniony przez prof. dr. hab. Hansa Hofmanna z Instytutu Statystyki i Ekonometrii Uniwersytetu w Hamburgu – pobrany został z repozytorium Uniwersytetu Kalifornijskiego¹, zawiera 1000 obserwacji (klientów banku) należących do jednej z 2 klas (klient „dobry” – udzielenie kredytu, klient „zły” – odmowa kredytu). Zadanie polega na zaklasyfikowaniu nowego klienta banku do jednej z tych grup. Obserwacje w zbiorze `german credit` charakteryzowane są przez 20 zmiennych objaśniających (13 nominalnych i 7 ilorazowych).

- X_1 – środki dostępne na ROR
- X_2 – okres użytkowania ROR-u
- X_3 – historia kredytowa
- X_4 – cel kredytu
- X_5 – wnioskowana kwota kredytu
- X_6 – oszczędności na lokatach
- X_7 – czas aktualnego zatrudnienia
- X_8 – rata kredytu jako procent przychodu

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

- X9* – płeć i stan cywilny
- X10* – żyranci
- X11* – czas ostatniego zameldowania
- X12* – majątek trwały/ruchomy
- X13* – wiek
- X14* – inne obciążenia spłatami pożyczek
- X15* – rodzaj mieszkania
- X16* – liczba kredytów w danym banku
- X17* – zatrudnienie
- X18* – liczba osób na utrzymaniu
- X19* – telefon
- X20* – pracownik zagraniczny
- Y* – klasa ryzyka (1 – udzielenie kredytu, 2 – odmowa kredytu)

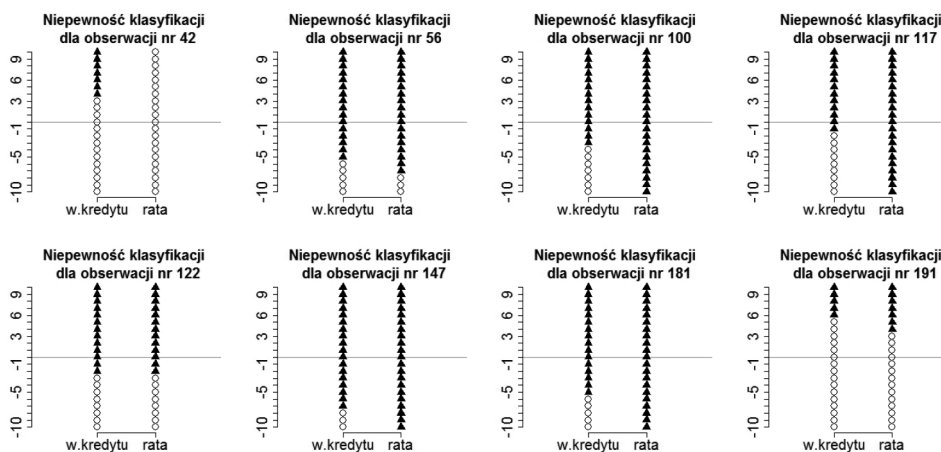
Ze względu na występowanie w zbiorze danych zmiennych nominalnych wielostanowych wystąpiła konieczność zakodowania tychże zmiennych za pomocą zmiennych sztucznych, co stanowiło już realizację pierwszego kroku procedury – wstępne przetwarzanie danych do analizy. W wyniku tego przetwarzania otrzymano zbiór uczący charakteryzowany przez 61 zmiennych objaśniających (54 zmienne sztuczne i 7 ilorazowych).

Wszystkie obliczenia zostały zrealizowane z wykorzystaniem programu statystycznego **R** i jego dodatkowych pakietów, które posłużyły do napisania autorskiego kodu realizującego przedstawioną procedurę.

W pierwszym kroku procedury opisanej w tab. 1 dodatkowo zredukowano wymiar przestrzeni danych metodą eliminacji zmiennych nieistotnych, wykorzystując technikę sekwencyjnej eliminacji z maksymalną wartością wskaźnika Randa jako kryterium eliminacji [Rand 1971; Trzęsiok 2009]. Liczba zmiennych objaśniających została zredukowana z 61 do 19. Następnie zbiór danych na potrzeby ilustracji działania procedury podzielono losowo na część uczącą (50%) i testową (50%). Część testowa posłużyła w dalszych krokach jako zbiór rozpoznawany, więc usunięto z niej zmienną wskazującą przynależność do klas.

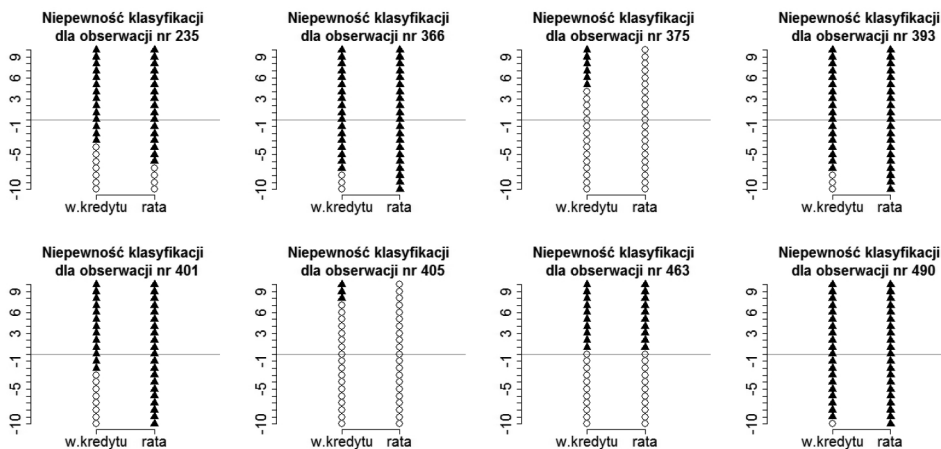
W drugim kroku na części uczącej zbudowano model klasyfikacyjny metodą wektorów nośnych (*SVM*) [Vapnik 1998] z wielomianową funkcją jądrową, przy czym *quasi*-optymalne wartości parametrów metody *SVM* dobrano symulacyjnie, wykorzystując sprawdzanie krzyżowe z podziałem zbioru uczącego na 10 segmentów [Trzęsiok 2004]. Wybór metody wektorów nośnych do zilustrowania procedury był arbitralny, lecz podyktowany własnościami tej metody: małą interpretowalnością modelu, na ogół wysoką dokładnością predykcji, nieliniowością postaci funkcji dyskryminującej [Trzęsiok 2008] – jest to więc metoda, dla której zwiększenie interpretowalności modelu jest zadaniem ważnym i stanowiącym spore wyzwanie.

W kroku trzecim zidentyfikowano dwie zmienne ilorazowe w zbiorze 19 zmiennych objaśniających użytych w modelowaniu: wnioskowana kwota kredytu oraz rata kredytu jako procent przychodu. Następnie w czwartym



Rys. 1. Wykres ilustrujący wyniki działania procedury wzbogacania metody klasyfikacji (tu *SVM*) o modelowanie niepewności. Wizualizacja dla pierwszych 8 z 16 obserwacji, które zidentyfikowano jako trudne do sklasyfikowania (klasyfikacja niestabilna)

Źródło: opracowanie własne.



Rys. 2. Wykres ilustrujący wyniki działania procedury wzbogacania metody klasyfikacji (tu *SVM*) o modelowanie niepewności. Wizualizacja dla drugiej grupy 8 z 16 obserwacji, które zidentyfikowano jako trudne do sklasyfikowania (klasyfikacja niestabilna)

Źródło: opracowanie własne.

kroku ustalono, że odchylenia do 10% (*in plus* oraz *in minus*) wartości oryginalnej będą uwzględniane w procedurze (w analizie wrażliwości). W kolejnym kroku poddano klasyfikacji wszystkie oryginalne obserwacje ze zbioru rozpoznawanego, wykorzystując model zbudowany w drugim kroku. W kroku szóstym dla każdej obserwacji ze zbioru rozpoznawanego utworzono zbiór zmodyfikowanych obserwacji z małymi (nie większymi niż 10%) odchyleniami wartości każdej zmiennej ilorazowej z osobna. Następnie zrealizowano krok siódmy i ósmy, tj. przeprowadzono klasyfikację nowo powstałych (w kroku 6.) obserwacji i sprawdzono, czy predykcja różni się od tej dla obserwacji pierwotnej. W przypadku pozytywnego wyniku weryfikacji informację uzupełniono szczególnym typem wykresu, który przedstawiono na rysunku 1 i 2.

W wyniku działania przedstawionej procedury zidentyfikowano 16 obserwacji (na 1000), których klasyfikacje uznano za niestabilne. Na rysunku 1 oraz 2 wypełnionymi trójkątami zaznaczono te obserwacje, dla których wskazanie modelu było negatywne (klient „zły” – wniosek kredytowy do odrzucenia), zaś okręgami zaznaczono te obserwacje, dla których model wskazał klasyfikację do grupy wniosków kredytowych do zaakceptowania. Na poziomie zera na osi rzędnych znajdują się obserwacje empiryczne z oryginalnymi wartościami zmiennych (zaznaczone linią poziomą jako punkty odniesienia), zaś powyżej i poniżej tego poziomu zaznaczono klasyfikacje dla odpowiednio zmodyfikowanych obserwacji. Szczególnie wyrazisty jest przypadek klienta (obserwacji) o numerze 463 (na rys. 2). Widać, że klient ten został sklasyfikowany przez model *SVM* jako klient „dobry”, czyli osoba, której należy udzielić kredytu, ale wystarczyłoby, że wnioskowana kwota kredytu byłaby o 1% większa, to model dawałby przeciwne wskazanie. Podobnie jest dla drugiej zmiennej ilorazowej: gdyby wyliczona rata kredytu jako procent przychodu była o 1% większa, to zgodnie z modelem wniosek kredytowy powinien zostać odrzucony. Świadczy to o tym, że ten klient znajduje się na brzegu klasy klientów „dobrych”, co wobec zakładanych błędów pomiaru zmiennych objaśniających oznacza duży poziom niepewności otrzymanej klasyfikacji dla tej obserwacji (wykryty przez analizę wrażliwości).

Innym wyrazistym przykładem ilustrującym wynik działania procedury jest przypadek obserwacji nr 117 (na rys. 1). Jest to klient, którego wniosek kredytowy został przez model wskazany jako „wniosek do odrzucenia”, ale wystarczyłoby, żeby wnioskowana kwota kredytu była o 2% mniejsza i wtedy klasyfikacja tego klienta zmieniałaby się.

4. Zakończenie

W artykule przedstawiono procedurę wzbogacającą metody klasyfikacji danych w zdolność do wskazywania, czy klasyfikacja danego obiektu jest trudna do rozstrzygnięcia w przypadkach, gdy jest to uzasadnione. Przedstawiona procedura wykorzystuje analizę wrażliwości i jest uniwersalna, tzn. może być stosowana z róż-

nymi metodami klasyfikacji. Dodatkowym walorem procedury jest zaproponowana w niej metoda wizualizacji wyników przeprowadzonej analizy wrażliwości. Ponadto do zalet zaproponowanej procedury zaliczyć należy kilka cech, które wynikają po części z założeń przyjętych przy jej konstruowaniu: procedura ma prostą i przejrzystą konstrukcję, nie zwiększa znacząco czasu obliczeniowego w porównaniu z użyciem samej metody klasyfikacji danych, daje łatwo interpretowalne rezultaty, pozwala na pozyskanie dodatkowej wiedzy o badanym zjawisku ze zbudowanego modelu, uzupełnia informacje o jakości modelu jako całości (mierzonej np. przez błąd klasyfikacji) o poziom zaufania do wyników predykcji dla pojedynczej obserwacji.

Wśród wad przedstawionego podejścia można wymienić to, że nie bada wrażliwości na interakcje małych zmian kilku zmiennych objaśniających (jest to konsekwencja przyjętej zasady *ceteris paribus*). Ponadto jako wadę należy wskazać subiektywizm w definiowaniu tego, co należy rozumieć jako „małe odchylenie” od oryginalnych wartości zmiennych objaśniających.

Literatura

- Bellman R.E., Kalaba R., Zadeh L.A., 1966, *Abstraction and pattern classification*, Journal of Mathematical Analysis and Applications, 13, s. 1-7.
- Hong D.H. i Choi C.-H., 2000, *Multicriteria fuzzy decision-making problems based on vague set theory*, Fuzzy Sets and Systems, 114(1), s. 103-113.
- Pawlak Z., 1982, *Rough sets*, International Journal of Computer and Information Sciences, 11, s. 341-356.
- Rand W.M., 1971, *Objective criteria for the evaluation of clustering methods*, Journal of the American Statistical Association, 66 (336), s. 846-850.
- Trzęsiok M., 2004, *Analiza wybranych własności metody dyskryminacji wykorzystującej wektory nośne*, [w:] Barczak A.S. (red.), *Postępy ekonometrii*, Wydawnictwo Akademii Ekonomicznej w Katowicach, Katowice, s. 331-342.
- Trzęsiok M., 2006, *Metoda wektorów nośnych na tle innych metod wielowymiarowej analizy danych*, [w:] Jajuga K., Walesiak M. (red.), *Klasyfikacja i analiza danych – teoria i zastosowania*, Taksonomia 13, Wydawnictwo Akademii Ekonomicznej we Wrocławiu, Wrocław, s. 536-542.
- Trzęsiok M., 2008, *Empiryczna ocena wrażliwości metody wektorów nośnych na występowanie obiektów błędnie sklasyfikowanych w zbiorze uczącym*, Studia Ekonomiczne, nr 50, s. 151-159.
- Trzęsiok M., 2009, *Problem doboru zmiennych do modelu dyskryminacyjnego budowanego metodą wektorów nośnych*, Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu, nr 47, Taksonomia 16, s. 214-222.
- Vapnik V., 1998, *Statistical Learning Theory*, John Wiley & Sons, N.Y.