

Dr inż. Maciej Huk, mgr inż. Michał Szczepanik

Instytut Informatyki
Politechnika Wrocławska
Ul. Wyb. Wyspiańskiego nr 27, 50-370 Wrocław, Polska
tel. (0-71) 320-35-16, fax. (0-71) 321-10-18
E-mail: Maciej.Huk@pwr.wroc.pl, Michal.Szczepanik@pwr.wroc.pl

Prawdopodobieństwo błędu klasyfikatorów złożonych dla problemów wieloklasowych

Słowa kluczowe: *klasyfikatory złożone, głosowanie większościowe, problemy wieloklasowe*

Streszczenie. W niniejszym artykule rozważamy systemy złożonych klasyfikatorów z głosowaniem większościowym dla przypadku problemów wieloklasowych, wykorzystujące wielowartościowe klasyfikatory bazowe. Stosując bezpośrednią reprezentację prawdopodobieństwa błędnej klasyfikacji dla analogicznych systemów w problemach dwuklasowych, otrzymujemy ogólny wzór na prawdopodobieństwo błędu klasyfikacji w przypadku wieloklasowym. Tym samym rozszerzamy teoretyczne analizy tego zagadnienia pierwotnie przeprowadzone dla problemów dwuklasowych przez Hansena i Salomona i ciągle wykorzystywane przez Kunchevę i innych badaczy. Pozwala nam to zaobserwować istotną zależność maksymalnego dopuszczalnego poziomu prawdopodobieństwa błędów klasyfikatorów bazowych od liczby rozważanych przez nie klas. Wskazuje to na możliwość poprawy parametrów klasyfikatorów złożonych dla problemów wieloklasowych, co może mieć niebagatelne znaczenie dla dalszych ich zastosowań w licznych dziedzinach nauki i przemysłu, z uwzględnieniem zagadnień diagnostyki maszyn oraz badania niezawodności systemów.

1. Wstęp

Systemy wielu współpracujących klasyfikatorów, znane także pod nazwą komitetów klasyfikatorów a w literaturze najczęściej nazywane klasyfikatorami złożonymi, były rozważane w wielu artykułach [5, 10, 13, 21, 23, 29, 34] oraz książkach [6, 8, 12, 18]. Podejścia tego typu, w ramach których budowanych jest wiele niezależnych hipotez a ich wyniki są łączone w celu zwiększenia dokładności klasyfikacji [3, 7] są często uznawane za jedno z głównych osiągnięć w dziedzinie indukcyjnego uczenia się w ostatniej dekadzie [2, 12, 19, 20, 27]. W efekcie, metodyka klasyfikatorów złożonych była z powodzeniem wykorzystywana do poprawy właściwości klasyfikacyjnych różnorodnych modeli w wielu dziedzinach takich jak finanse i bankowość [22], bioinformatyka [32], medycyna [24], inżynieria produkcji [28], geografia [4], bezpieczeństwo informacji [16, 25], odzyskiwanie danych [10] oraz w systemach doradczych [17]. Na jej bazie zaproponowano także wiele rozwiązań problemów diagnostyki maszyn i systemów elektronicznych [31, 35] oraz badania niezawodności systemów [14, 30]. Rozwiązania tego typu mogą stanowić cenne uzupełnienie innych, dotąd wykorzystywanych podejść [26, 33, 36].

W niniejszym artykule rozszerzamy teoretyczną analizę prawdopodobieństwa błędu klasyfikatorów złożonych pierwotnie przeprowadzoną dla problemów dwuklasowych przez Hassena i Salomona [15] i nadal wykorzystywaną przez Kunchevę i innych badaczy [18-20, 29]. Rozważamy ogólny przypadek problemów wieloklasowych i komitety klasyfikatorów bazowych wykorzystujące klasyczne głosowanie większościowe. Na tej bazie wyprowadzamy ogólny wzór na prawdopodobieństwo błędu klasyfikatora złożonego w

problemach z liczbą klas większą niż dwa oraz dla dowolnej liczby klasyfikatorów bazowych z jednakowymi wartościami prawdopodobieństwa błędu a posteriori. W trakcie powyższego przedstawiamy również, co jest często pomijane, jak zmienia się powszechnie znany wzór na prawdopodobieństwa błędu klasyfikatora złożonego dla problemów dwuklasowych, gdy liczba klasyfikatorów bazowych nie jest ograniczona do liczb nieparzystych. Analiza otrzymanych rezultatów wskazuje możliwość wykorzystania wielowartościowych klasyfikatorów bazowych do poprawy parametrów klasyfikatorów złożonych, nawet dla bardzo trudnych problemów klasyfikacji.

2. Prawdopodobieństwo błędu klasyfikatorów złożonych dla problemów dwuklasowych

Niech $D = \{D_1, \dots, D_L\}$ będzie zbiorem L klasyfikatorów $D_i: \mathfrak{X}^n \rightarrow \Omega$, gdzie $\Omega = \{\omega_1, \dots, \omega_C\}$, przypisujących etykietę klasy $\omega_j \in \Omega$ do wektora danych wejściowych $\mathbf{x} \in \mathfrak{X}^n$. Zakłada się, że klasyfikatory ze zbioru D mogą być z powodzeniem wykorzystane do budowy komitetu klasyfikatorów, jeśli ich błędy są wzajemnie nieskorelowane lub negatywnie skorelowane [1] oraz gdy dla każdego klasyfikatora bazowego D_i jego prawdopodobieństwo błędu a posteriori P_S^i jest mniejsze niż 0.5. W przypadku problemów dwuklasowych ($K=2$) z wykorzystaniem głosowania większościowego sytuacja jest stosunkowo prosta i prawdopodobieństwo błędu P_E klasyfikatora złożonego jest często wyrażane jako:

$$P_E = \sum_{j=j_0}^L \binom{L}{j} P_S^j [1 - P_S]^{L-j}, \quad (1)$$

gdzie L jest nieparzyste, wszystkie klasyfikatory mają równe prawdopodobieństwa błędu a posteriori P_S oraz początkowa wartość j_0 jest najmniejszą liczbą klasyfikatorów bazowych dających niepoprawny wynik, która prowadzi do błędnej decyzji komitetu klasyfikatorów.

Należy jednak pamiętać, że dla problemów wieloklasowych ograniczenie liczby klasyfikatorów bazowych L do wartości nieparzystych nie eliminuje możliwości tego, że klasyfikatory bazowe zremisują. W takim zaś przypadku zazwyczaj klasa jest ostatecznie wybierana losowo spośród klas remisujących, które uzyskały największą liczbę głosów, przy czym żadna z nich nie jest faworyzowana. Uwzględniając powyższe, czynnik prawdopodobieństwa błędu klasyfikatora złożonego związany z remisami nie może być pominięty. W efekcie, poszukując wskazań do dalszej analizy problemów wieloklasowych, możemy pominąć założenie o nieparzystości liczby klasyfikatorów bazowych L i rozszerzyć wyrażenie (1) do postaci:

$$P_E = \sum_{j=j_0}^L \binom{L}{j} P_S^j [1 - P_S]^{L-j} + \frac{1}{2} \delta(L \bmod 2, 0) \binom{L}{L/2} P_S^{L/2} [1 - P_S]^{L/2}, \quad (2)$$

gdzie

$$j_0 = \begin{cases} (L+1)/2 & : L \bmod 2 > 0 \\ L/2 + 1 & : L \bmod 2 = 0 \end{cases}, \quad (3)$$

a $\delta(x, y)$ jest deltą Kronekera:

$$\delta(x, y) = \begin{cases} 1 & : x = y \\ 0 & : x \neq y \end{cases}. \quad (4)$$

Współczynnik $\frac{1}{2}$ przed deltą Kronekera w wyrażeniu (2) to prawdopodobieństwo losowego wyboru niepoprawnej klasy podczas remisu klasyfikatorów bazowych, a występujący po nim

symbol Newtona $\binom{L}{\frac{L}{2}}$ określa liczbę możliwych remisów pomiędzy klasyfikatorami bazowymi dla problemu dwuklasowego, gdy L jest parzyste.

3. Prawdopodobieństwo błędu klasyfikatorów złożonych dla problemów wieloklasowych

Pierwszym krokiem do znalezienia ogólnego równania dla prawdopodobieństwa błędu klasyfikatora złożonego dla problemów wieloklasowych jest przekształcenie wyrażenia (2) tak by każda składowa sumarycznego prawdopodobieństwa była jawnie związana z liczbami głosów oddawanymi przez klasyfikatory bazowe na poszczególne klasy. Ponieważ bez utraty ogólności rozważań możemy przyjąć, iż prawidłową odpowiedzią klasyfikatora złożonego jest klasa o indeksie 1, to dzięki prostym przekształceniom algebraicznym możemy zauważyć, że prawa strona wyrażenia (1) może przyjąć postać:

$$\sum_{k_1=0}^L \sum_{k_2=0}^L \left(\binom{L}{k_2} P_S^{k_2} (1-P_S)^{L-k_2} \cdot \delta(k_1+k_2, L) H(k_2-k_1) \right), \quad (5)$$

gdzie k_1 i k_2 reprezentują liczby głosów oddanych przez L klasyfikatorów bazowych odpowiednio na klasę 1 i klasę 2. Wprowadzona delta Kroneckera gwarantuje przy tym, że tylko te kombinacje głosów są brane pod uwagę, dla których suma głosów dla wszystkich klas równa jest liczbie klasyfikatorów bazowych:

$$k_1 + k_2 = L, \quad (6)$$

zaś H to funkcja Heaviside'a służąca do wyboru czynników dla których $k_2 > k_1$:

$$H(x) = \begin{cases} 1 & : x > 0 \\ 0 & : x \leq 0 \end{cases}. \quad (7)$$

Wreszcie, dzięki dalszemu wykorzystaniu (6) do obliczenia $L - k_2$, oraz poprzez wprowadzenie oznaczeń:

$$P_1 = 1 - P_S \text{ i } P_2 = P_S \quad (8)$$

na prawdopodobieństwa głosowania odpowiednio na klasę 1 i 2, możemy przepisać (5) w formie:

$$\sum_{k_1=0}^L \sum_{k_2=0}^L \left(\frac{L!}{k_1! k_2!} P_1^{k_1} P_2^{k_2} \delta(k_1+k_2, L) H(k_2-k_1) \right). \quad (9)$$

Analogicznie, prawy człon prawej strony wyrażenia (2) może być przekształcony w:

$$\sum_{k_2=0}^L \left(\frac{1}{2} \binom{L}{k_2} P_S^{k_2} [1-P_S]^{L-k_2} \delta(k_2, \frac{L}{2}) \right). \quad (10)$$

Następnie, ze względu na to iż w przypadku remisu mamy $k_1 = k_2 = \frac{L}{2}$, formuła (10) może być zapisana jako:

$$\sum_{k_1=0}^L \sum_{k_2=0}^L \left(\frac{1}{2} \delta(k_1+k_2, L) \delta(k_1, k_2) \frac{L!}{k_1! k_2!} P_1^{k_1} P_2^{k_2} \right). \quad (11)$$

W rezultacie, po połączeniu (5) i (11) oraz uporządkowaniu, wzór na prawdopodobieństwo błędu klasyfikatora złożonego dla problemu dwuklasowego (2) może być przedstawiony jako:

$$P_E = \sum_{k_1=0}^L \sum_{k_2=0}^L \left(\delta(k_1+k_2, L) \left[H(s_2-s_1) + \frac{1}{2} \delta(k_1, k_2) \right] \frac{L!}{k_1! k_2!} P_1^{k_1} P_2^{k_2} \right). \quad (12)$$

Wyrażenie (12) ukazuje naturalną metodę wyznaczania wartości prawdopodobieństwa błędu klasyfikatora złożonego dla przypadku wieloklasowego ($K > 2$) – poprzez dodanie kolejnych sumowań przypisanych kolejnym klasom rozważanym przez klasyfikatory bazowe. Łatwo zauważyć, że w takim przypadku dodatkowej analizy wymaga tylko część wyrażenia (12) znajdująca się w nawiasach kwadratowych. Funkcja Heaviside'a dostarcza w jej przypadku informację czy klasa poprawna otrzymała mniej głosów niż inne klasy. Tym samym dla wielu klas powinna być ona zastąpiona przez formułę:

$$H_E = H\left(\sum_{i=2}^K H(s_i - s_1)\right), \quad (13)$$

która przyjmuje wartość jeden, gdy jedna lub więcej klas otrzymała od klasyfikatorów bazowych więcej głosów niż klasa poprawna, oraz wartość zero w pozostałych przypadkach. Druga, prawa część pod nawiasem kwadratowym w (12) - delta Kroneckera - może być identyfikowana jako element określający liczbę klas remisujących z klasą poprawną, dodatkowo pomnożoną przez prawdopodobieństwo wylosowania niepoprawnej klasy podczas remisu. W ogólnym przypadku ($K > 2$) liczba remisów może być wyznaczona za pomocą wzoru:

$$H_D = \sum_{i=2}^K \delta(s_1, s_i), \quad (14)$$

dzięki czemu prawdopodobieństwo wylosowania niepoprawnej klasy podczas remisu jest dane przez wyrażenie:

$$\frac{H_D}{H_D + 1}. \quad (15)$$

Uwzględniając powyższe łatwo obliczyć, że prawdopodobieństwo błędu klasyfikatora złożonego dla problemów wieloklasowych jest równe:

$$P_E = \sum_{k_1=0}^L \sum_{k_2=0}^L \cdots \sum_{k_K=0}^L \left(\delta(k_1 + k_2, L) \left[H(s_2 - s_1) + \frac{1}{2} \delta(k_1, k_2) \right] \frac{L!}{k_1! k_2!} P_1^{k_1} P_2^{k_2} \right), \quad (16)$$

gdzie suma prawdopodobieństw oddania głosów na poszczególne klasy:

$$\sum_{i=1}^K P_i = 1. \quad (17)$$

Jednocześnie jednak warto zauważyć, że czynnik

$$L! \prod_{i=1}^K \frac{1}{k_i!} \quad (18)$$

jest wielomianowym współczynnikiem P_{MF} wielomianowego rozkładu prawdopodobieństwa, więc wyrażenie (16) może być zapisane ostatecznie jako:

$$P_E = \sum_{k_1=0}^L \sum_{k_2=0}^L \cdots \sum_{k_K=0}^L \left(P_{MF} \left[H_E + (1 - H_E) \left(1 - \frac{1}{1 + H_D} \right) \right] \right), \quad (19)$$

gdzie:

$$P_{MF} = f(k_1, k_2, \dots, k_K, L, P_1, P_2, \dots, P_K) \quad (20)$$

jest funkcją rozkładu prawdopodobieństwa rozkładu wielomianowego dla nieujemnych liczb całkowitych k_1, k_2, \dots, k_K .

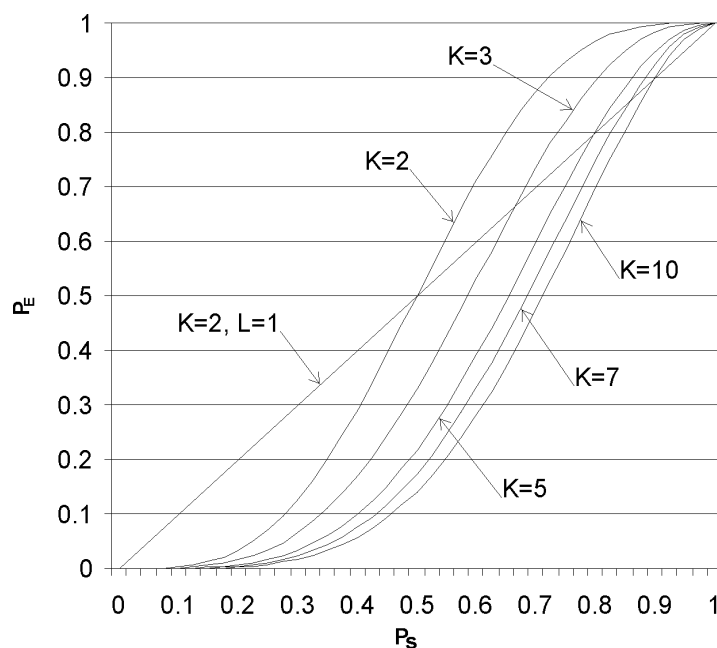
4. Symulacje i dyskusja wyników

Wprowadzone w poprzednim rozdziale wyrażenie (19), zostało w pierwszej kolejności zweryfikowane doświadczalnie z wykorzystaniem symulacji statystycznych

systemu z wieloma klasyfikatorami. Ze względu na wysoki koszt obliczeniowy tego typu symulacji, pod uwagę były brane tylko przypadki liczności klas K od 2 do 10, liczby klasyfikatorów bazowych L od 1 do 100 i wybrane wartości prawdopodobieństw błędnych klasyfikacji klasyfikatorów bazowych P_S (0; 0,1; 0,3; 0,5; 0,7; 0,9 i 1). Podczas symulacji dla każdego zestawu parametrów przeprowadzono 10^6 głosowań, w których odpowiedzi poszczególnych klasyfikatorów bazowych generowano losowo, z zastosowaniem standardowego generatora liczb losowych zawartego w Borland Object Pascal w bibliotece System.

Uzyskane wyniki wykazały wysoką spójność rezultatów przeprowadzonych symulacji i wartości wyliczonych na podstawie wzoru (19). Dla wszystkich rozważanych wartości parametrów wyniki symulacji różniły się od obliczonego prawdopodobieństwa błędu o nie więcej niż 2,7% wartości wyliczanych (średnio o 0,043%). Dodatkowo, dla przypadku problemów dwuklasowych obie metody dały wyniki zgodne z wartościami wyrażenia (2).

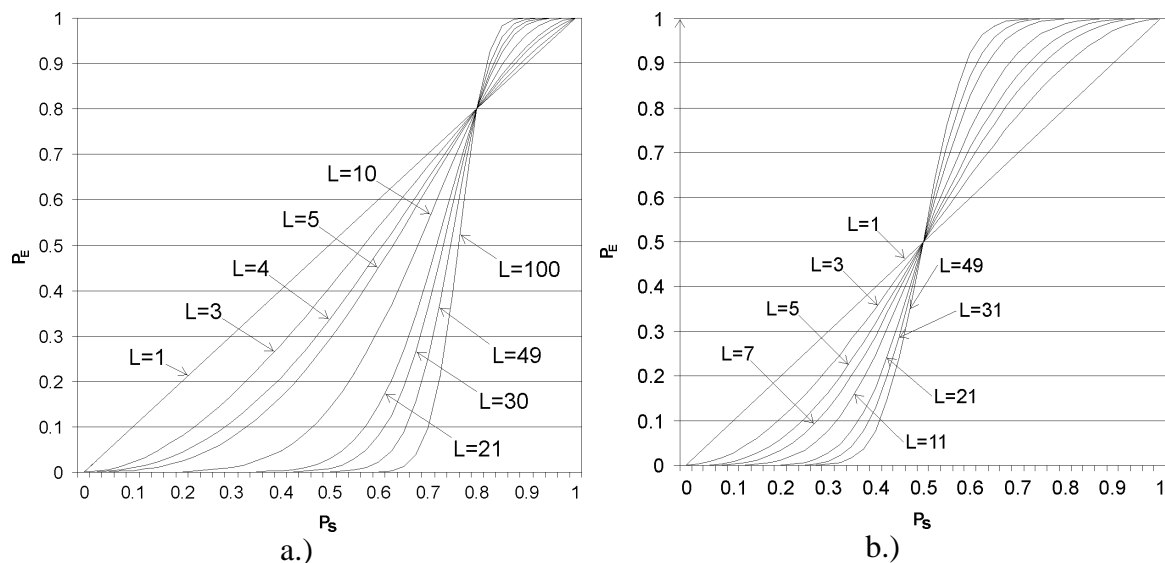
Na podstawie powyższych badań obserwowaliśmy, jak prawdopodobieństwo błędu klasyfikatora złożonego zmienia się wraz ze wzrostem liczby klas (patrz rys. 1). Dla typowego przykładu $L = 21$ i $P_S = 0,3$ dla dwóch klas prawdopodobieństwem błędu wynosi $P_E \approx 0,0264$, ale dla trzech i pięciu klas wynosi ono odpowiednio już tylko 0,00202 oraz 0,000126. Jest to rezultat wzrastającej liczby klas innych niż prawidłowa. Błędne głosy są rozproszone na wszystkie $K - 1$ niepoprawne klasy i w efekcie średnia łączna liczba głosów dla poszczególnych niepoprawnych klas maleje wraz ze wzrostem liczby klas K , co nie dotyczy klasy poprawnej.



Rys. 1. Prawdopodobieństwo błędu klasyfikatora złożonego P_E w funkcji prawdopodobieństwa błędu P_S siedmiu klasyfikatorów bazowych ($L=7$), z negatywnie skorelowanymi wzajemnymi błędami, dla różnych licznosci klas K .

Bardzo ciekawe jest również to, że dla liczby klas K większej niż 2, górny limit prawdopodobieństwa a posteriori klasyfikatorów bazowych, przy którym możliwe jest budowanie skutecznego klasyfikatora złożonego, jest większy niż 0,5 (porównaj rys. 2a i rys. 2b). W obliczu praktycznych trudności w tworzeniu dużych zbiorów klasyfikatorów bazowych o niskim prawdopodobieństwie błędów i jednocześnie o wysokim stopniu braku korelacji pomiędzy popełnianymi przez nie błędami, obserwowany rezultat sugeruje

możliwość łatwiejszego budowania klasyfikatorów złożonych dla problemów wieloklasowych przez dopuszczenie do rozważań także klasyfikatorów bazowych popełniających błędy częściej niż w połowie przypadków.



Rys. 2. Prawdopodobieństwo błęd klasyfikatora złożonego P_E w funkcji prawdopodobieństwa błęd P_S dla różnych liczb L klasyfikatorów bazowych, z negatywnie skorelowanymi wzajemnymi błędami, dla pięciu a.) i dwóch b.) klas.

Dla przykładu - gdy liczba klas $K = 5$ i liczba klasyfikatorów bazowych $L = 21$, prawdopodobieństwo błęd klasyfikatorów bazowych $P_S = 0,6$ skutkuje prawdopodobieństwem błęd klasyfikatora złożonego na poziomie $P_E \approx 0,146$, co jest wartością lepszą niż dla zgadywania losowego. Dodatkowo, poprzez zwiększenie liczby klasyfikatorów bazowych do 100, powyższe prawdopodobieństwo błęd klasyfikatora złożonego można zmniejszyć do zaledwie 0,000815. Należy jednak pamiętać, że prezentowane wyniki zostały uzyskane przy założeniu, że wzajemne błędy klasyfikatorów bazowych są w pełni nieskorelowane lub negatywnie skorelowane, co jest trudne do osiągnięcia w praktyce. Częściowa korelacja błędów może powodować zmiany poszczególnych wartości podanych wyżej prawdopodobieństw, jednak nie powinna wpływać na charakter uzyskiwanych rezultatów.

5. Podsumowanie i plany badawcze

W niniejszej pracy formalnie przedstawiono wzór na prawdopodobieństwo błęd klasyfikatora złożonego dla problemów wieloklasowych. Jego szczegółowe wyprowadzenie zostało oparte na powszechnie znanym, analogicznym wzorze dla problemów dwuklasowych, który został w tej pracy dodatkowo rozszerzony na przypadki z parzystą liczbą klasyfikatorów bazowych.

Symulacyjna analiza uzyskanych wzorów ukazała, że wzrost liczby klas rozważanych przez klasyfikatory bazowe powyżej dwóch pozwala obniżyć prawdopodobieństwo błęd klasyfikatora złożonego bez zwiększania liczebności komitetu. Interesujące jest również przy tym to, że przy założeniu braku lub negatywnej korelacji wzajemnych błędów klasyfikatorów bazowych, górny limit ich prawdopodobieństwa błęd P_S , przy którym zasadne jeszcze jest budowanie klasyfikatora złożonego, wzrasta wraz ze wzrostem liczby rozważanych klas.

W rezultacie powyższego, przejście od schematu hipotez dwuwartościowych do wielowartościowych, ułatwia tworzenie dużych zbiorów różnorodnych klasyfikatorów bazowych, a co za tym idzie - także doskonalszych klasyfikatorów złożonych. Może mieć to

niebagatelne znaczenie dla dalszych zastosowań tego typu metod w licznych dziedzinach nauki i przemysłu – w tym w zagadnieniach utrzymania i diagnostyki maszyn oraz badania niezawodności systemów.

W dalszych pracach będziemy badać, w jaki sposób częściowa korelacja między błędami wielowartościowych klasyfikatorów bazowych zmienia prawdopodobieństwa popełniania błędów klasyfikatorów złożonych dla liczb klas większych niż 2. Będziemy także opracowywać efektywne obliczeniowo wyrażenia do estymacji wartości wyprowadzonej formuły dla liczb klas rzędu stu i większych.

Literatura

1. Ali K, Pazzani M. Error reduction through learning multiple descriptions. *Machine Learning* 1996; 24(3): 173-206.
2. Bian S, Wang W. On diversity and accuracy of homogeneous and heterogeneous ensembles. *IOS Press Amsterdam*: 2007, 4(2): 103-128.
3. Brown G, Wyatt J, Harris R, Yao X. Diversity creation methods: A survey and categorization. *Journal of Information Fusion* 2005; 6(1).
4. Bruzzone L, Cossu R, Vernazza G. Detection of land-cover transitions by combining multivariate classifiers. *IOS Press Amsterdam*: 2007, 25(13): 1491-1500.
5. Buhlmann P, Hothorn T. Boosting algorithms: Regularization, Prediction and Model Fitting. *Statistical Science* 2007; 22(4): 477-505.
6. Claeskens G, Hjort N. *Model Selection and Model Averaging*. Volume 27 of Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press 2008.
7. Dietterich T. An experimental comparison of three methods for constructing ensembles of decision trees: Bagging, boosting, and randomization., *Machine Learning* 2000; 40(2): 139-157.
8. Dietterich T. (Ensemble learning, in M. ARBIB (Ed.)) *The Handbook of Brain Theory and Neural Networks*. Second ed., Cambridge: 2002, 405-408.
9. Elovici, Y, Shapira B, Kantor P. A decision theoretic approach to combining information filters: Analytical and empirical evaluation., *Journal of the American Society for Information Science and Technology* 2006; 57(3): 306-320.
10. Evgeniou T, Pontil M, Elisseeff A. Leave one out error, stability, and generalization of voting combinations of classifiers, *Machine Learning* 2004; 55(1): 71-97.
11. Freund Y, Lu J, Schapire R. *Boosting: Models, Applications and Extensions*., Chapman and Hall/CRC 2010.

12. Freund Y, Schapire R. Experiments with a new boosting algorithm., *Machine Learning: Proceedings of the Thirteenth International Conference (ICML 96)*. SanFrancisco: 1996, 148-156.
13. Fumera G, Roli F. A theoretical and experimental analysis of linear combiners for multiple classifier systems., *IEEE Transactions on Pattern Analysis and Machine Intelligence* 2005; 27(6): 942-956.
14. Halbiniak Z, Józwiak I.: Deterministic chaos in the processor load. *Chaos, Solitons and Fractals* 2007; 31(2): 409-416.
15. Hansen L, Salamon P. Neural network ensembles, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 1990; 12(10): 993-1001.
16. Jacak J, Józwiak I, Jacak L.: Composite fermions in braid group terms. *Open Systems and Information Dynamics* 2010; 17(1): 53-71.
17. Jahrer M, Tscher A, Legenstein R. Combining predictions for accurate recommender systems, *Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining*, Washington, DC, USA: 2010, 693-702.
18. Kuncheva L. *Combining Pattern Classifiers. Methods and Algorithms.*, Wiley 2004
19. Kuncheva L, Bezdek J, Sutton M. On combining multiple classifiers by fuzzy templates, *Proceedings of the on combining multiple classifiers by fuzzy templates. Conference*, Pensacola, Florida USA: 1990, 193-197.
20. Kuncheva L, Whitaker C, Shipp C, Duin R. Limits on the majority vote accuracy in classifier fusion, *Pattern Analysis and Applications* 2003; 6: 22-31.
21. Leapa N, Clemansa P, Bauer K, Oxley M. An investigation of the effects of correlation and autocorrelation on classifier fusion and optimal classifier ensembles, *International Journal of General Systems* 2008; 37(4): 475-498.
22. Leigh W, Purvis R, Ragusa J. Forecasting the nyse composite index with technical analysis, pattern recognizer, neural networks, and genetic algorithm: a case study in romantic decision support. *Decision Support Systems* 2002; 32(4): 361-377.
23. Liu Y, Yao X, Higuchi T. Evolutionary ensembles with negative correlation learning. *IEEE Transactions on Evolutionary Computation* 2000; 4(4): 380-387.
24. Mangiameli P, West D, Rampal R. Model selection for medical diagnosis decision support systems, *Decision Support Systems* 2004; 36(3): 247-259.
25. Menahem E, Shabtai A, Rokach L, Elovici Y. Improving malware detection by applying multi-inducer ensemble. *Computational Statistics and Data Analysis* 2009; 53(4): 1483-1494.

26. Niewczas A, Pieniak D, Bachanek T, Surowska B, Bieniaś J, Pałka K. Prognosing of functional degradation of bio-mechanical systems exemplified by the tooth-composite filling system. *Eksploatacja i Niezawodność - Maintenance and Reliability* 2010; 45(1): 23-34.
27. Opitz D, Shavlik J. Generating accurate and diverse members of a neural-network ensemble, *Advances in Neural Information Processing Systems*. MIT Press, Denver: 1996, 535-543.
28. Rokach L. Mining manufacturing data using genetic algorithm-based feature set decomposition. *International Journal of Intelligent Systems Technologies and Applications* 2008; 4(1): 57-78.
29. Rokach L. Taxonomy for characterizing ensemble methods in classification tasks: a review and annotated bibliography. *Computational Statistics and Data Analysis* 2009; 53(12): 4046-4072.
30. Santhanam P, Bassin K. Managing the maintenance of ported, outsourced, and legacy software via orthogonal defect classification. In *proc. IEEE International Conference on Software Maintenance* 2001; 726-734
31. Shahrtash S, Jamehbozorg A. A Decision-Tree-Based Method for Fault Classification in Single-Circuit Transmission Lines. *IEEE Transactions on Power Delivery* 2010; 25(4): 2190-2196.
32. Tan A, Gilbert D, Deville Y. Multi-class protein fold classification using a new ensemble machine learning approach. *Genome Informatics* 2003; 14: 206-217.
33. Tao J, Zhang Y, Chen X, Ming Z. Bayesian reliability growth model based on dynamic distribution parameters. *Eksploatacja i Niezawodność - Maintenance and Reliability* 2010; 46(2): 13-16.
34. Valentini G, Masulli F. Ensembles of learning machines. in M. M. and T. R. (Eds), *Neural Nets: 13th Italian Workshop on Neural Nets*, Vol. 2486 of *Lecture Notes in Computer Science*, Springer, Berlin: 2002, 3-19.
35. Xu D, Wu M, An J. Design of an expert system based on neural network ensembles for missile fault diagnosis. In *Proc. IEEE International Conference on Robotics, Intelligent Systems and Signal Processing* 2003, 2: 903-908.
36. Yu T, Cui W, Song B, Wang S. Reliability growth estimation for unmanned aerial vehicle during flight-testing phases. *Eksploatacja i Niezawodność - Maintenance and Reliability* 2010; 46(2): 43-47.