

Maciej Wasiak

Akademia Ekonomiczna we Wrocławiu

PRZYSPIESZENIE DZIAŁANIA ALGORYTMU GENETYCZNEGO W UCZENIU SYSTEMÓW ROZMYTYCH

1. Wstęp

Tempo automatycznego uczenia systemów wnioskujących jest jednym z kryteriów oceny opracowywanych i stosowanych algorytmów. W uczeniu systemów rozmytych stosowanych jest kilka metod, spośród których bardzo dobre efekty daje algorytm genetyczny. Niestety, słabym punktem systemów rozmytych jest stosunkowo długotrwały proces wnioskowania, co w przypadku metod uczenia wymagających przeprowadzenia nawet kilku milionów procesów wnioskujących przekłada się na ogólny długi okres uczenia.

Niezależnie od stanu zaawansowania aktualnej technologii, zadania zawsze wyprzedzają możliwości obliczeniowe. Dlatego też nieustannie trwają prace nad poprawą działania algorytmów uczących. Należy wspomnieć, że badacze publikujący uzyskane wyniki trafności wnioskowania, np. na typowych zbiorach testowych, dość rzadko podają informację o czasie oraz liczbie prób, jakich potrzebowali na osiągnięcie podanego wyniku. W niniejszym artykule opisano kilka prób poprawy efektywności działania algorytmu genetycznego zastosowanego do uczenia systemu wnioskującego wraz z dokładną charakterystyką przebiegu procesu uczenia. Ze względu na charakter modyfikacji, autor żywi nadzieję, że przedstawione próby mogą być użyteczne również dla uczenia systemów wnioskujących opartych na innych reprezentacjach wiedzy, a opisane niepowodzenia pozwolą innym uniknąć tych samych błędów i, być może, stanowić będą inspirację do dalszych usprawnień działania algorytmu genetycznego.

2. Opis systemu rozmytego i algorytmu genetycznego

Opisane poniżej próby poprawy procesu uczenia dotyczą systemu rozmytego, mającego za zadanie ustalić kilka podstawowych parametrów:

- wybór kluczowych zmiennych,
- określenie granic zbiorów rozmytych,
- określenie konkluzji dla reguł.

Zbiory rozmyte były z założenia niesprzeczne i kompletne, podobnie jak baza reguł była zawsze kompletna i nie zawierała reguł sprzecznych. Do wnioskowania rozmytego został użyty model Mamdaniego, agregacja według zasady *max-min* i taki sam zastosowano typ inferencji. Do opisanie zbiorów rozmytych zastosowano funkcje trapezowe, dopełniające się w ramach jednej zmiennej do jedności. Sposób wyostrzania przeprowadzano metodą środka ciężkości.

Parametrem zakładanym z góry był typ zbiorów rozmytych (w tym przypadku trapezowy) oraz liczba kluczowych zmiennych objaśniających. Na potrzeby opisywanych badań przyjęto wersję systemu, który miał samodzielnie wybrać dwie najlepsze zmienne spośród liczby posiadanych. Każda z wybranych zmiennych opisana była dwoma zbiorami rozmytymi, których granice ustalić miał AG. Razem daje to prostą bazę wiedzy złożoną z zaledwie czterech reguł. W trakcie projektowania systemu, okazało się, że możliwość swobodnego wyboru zmiennych używanych do wnioskowania najbardziej wpływa na wydłużenie czasu uczenia systemu przez AG.

We wczesnych etapach pracy nad systemem wyniki nie były satysfakcjonujące, nawet uzyskiwane na prostych zbiorach danych, takich jak IRIS, czy Wisconsin Breast Cancer. Kluczem do osiągnięcia akceptowalnej efektywności algorytmu uczącego było odpowiednie zakodowanie parametrów systemu rozmytego do postaci chromosomu. Ostatecznie, obok typowych elementów AG, zastosowano:

- kodowanie rzeczywiste,
- krzyżowanie konwencjonalne z prawdopodobieństwem 0,4,
- krzyżowanie uśredniające z prawdopodobieństwem 0,4,
- *steady-state*,
- mutacja z prawdopodobieństwem 0,02 i skalą mutacji 100%.

Niestety, o ile trafność wnioskowania była bardzo dobra: na zbiorze IRIS ok. 98%, a na zbiorze Wisconsin Breast Cancer 97,5%, o tyle czas uczenia pozostawiał wiele do życzenia, szczególnie przy zbiorach bardziej złożonych. Najszersze badania dotyczyły zastosowania wyżej opisanego systemu do oceny kondycji finansowej przedsiębiorstw (por. [14]). Baza przypadków liczyła ponad 300 zestawów po 26 wskaźników finansowych, spośród których system miał wybrać kilka najistotniejszych. Optymalizacja systemu tego rodzaju trwała od kilku do kilkunastu godzin.

Wiele zadań jest znacznie bardziej złożonych, stąd próby przyspieszenia procesu automatycznego uczenia. Poniżej przedstawiono kilka spośród wielu prób przyspieszenia pracy AG podejmowanych przez autora. Dodajmy, że próby te w większości zakończone były fiaskiem.

3. Przyspieszenie metodą wzrostu liczby przypadków

Jedna z testowanych koncepcji zakładała uczenie w początkowej fazie pracy AG z użyciem niewielkiej liczby przypadków – np. pierwszych 10%. Następnie stopniowo zwiększano liczbę przypadków aż do pełnej ich liczby. Za użyciem takiej strategii przemawiało przypuszczenie, że skuteczne będzie zaoszczędzenie czasu i poświęcenie go na większą liczbę generacji ewolucyjnych przy użyciu mniejszej liczby przypadków niż wszystkie posiadane. Przeprowadzono wiele testów na zbiorach IRIS oraz Wisconsin Breast Cancer, z których przykładowe przedstawiono w tab. 1.

Tabela 1. Przykłady rozwiązań przy początkowym ograniczaniu zbioru przypadków

Początkowa liczba przypadków	Początek zwiększania liczby przypadków	Tempo wzrostu liczby przypadków
10%	od pierwszej generacji	1 przypadek na 1 generację
10%	od pierwszej generacji	2 przypadki na 1 generację
10%	od 10 generacji	1 przypadek na 1 generację
10%	od 10 generacji	2 przypadki na 1 generację
20%	od pierwszej generacji	1 przypadek na 1 generację
20%	od pierwszej generacji	2 przypadki na 1 generację
20%	od 10 generacji	1 przypadek na 1 generację
20%	od 10 generacji	2 przypadki na 1 generację

Źródło: badania własne.

Testowano również wiele innych kombinacji – z początkową liczbą przypadków do 50%, z różnymi momentami i tempem zwiększania liczby przypadków. W części badań uzależniono wzrost liczby przypadków od postępów systemu w trafności wnioskowania. Badania wielokrotnie powtarzano w celu uzyskania wiarygodnych wyników.

Niestety nie odnotowano istotnej poprawy działania AG, a w większości badań wyniki były zdecydowanie gorsze niż przy uczeniu od początku całym zbiorem posiadanych przypadków.

4. Przyspieszenie metodą wzrostu liczby przypadków ze zmiennym startem

Obserwacja opisanych wyżej prób skłoniła do wniosku, że ograniczanie przypadków do zawsze takich samych powoduje bardzo wczesne dopasowanie się

uczonego systemu do tych właśnie danych. Później bardzo trudno jest to dopasowanie zmienić, stąd słabe wyniki uczenia. Zmodyfikowano więc proces tak, aby stopniowo w trakcie uczenia „przesuwał” ograniczoną próbkę danych, np. na początku przypadki od 1 do 20, potem od 2 do 21 itd. Jednocześnie stosowano stopniowe zwiększanie liczby przypadków.

Różne eksperymenty z zastosowaniem powyższej strategii dały lepsze rezultaty niż bez „przesuwania przypadków”, ale cały czas gorsze od całkowicie standardowego uczenia na podstawie całej posiadanej próbki.

5. Przyspieszenie metodą wstępnego uczenia z losowo dobieranymi przypadkami

Jedną z kolejnych prób skrócenia czasu uczenia systemu wnioskującego było ograniczanie początkowej ilości danych przy losowym doborze przypadków. Podobnie jak w opisanych wcześniej próbach, początkowa próbka wahała się w różnych eksperymentach od 10 do 50%. Losowanie było dokonywane bez zwracania, tak więc poszczególne przypadki były używane w procesie uczenia bez powtórzeń. W praktyce dla każdej generacji tworzony był nowy, niepełny zbiór danych z losowo dobranymi przypadkami.

Uzyskane efekty dla niektórych kombinacji początkowej liczby przypadków oraz tempa powiększania ich zbioru nie były wyraźnie gorsze od badania standardowego (z kompletną liczbą przypadków). Autor przypuszcza, że odpowiedni układ obu tych parametrów, wraz z dostosowaniem parametrów mutacji oraz krzyżowania, mógłby przynieść pozytywne efekty. Badania w tym zakresie są kontynuowane.

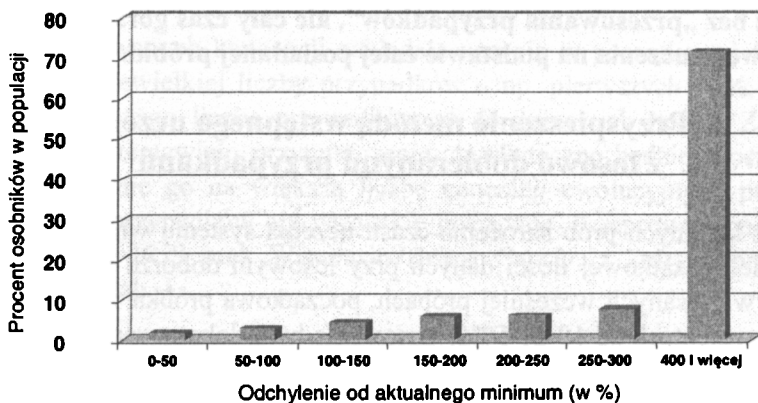
6. Przyspieszenie metodą wczesnej eliminacji chromosomów

Ostatnia z opisanych prób powstała na bazie obserwacji rozkładu jakości chromosomów w populacji po operacjach krzyżowania oraz mutacji. Okazało się, że typowym zjawiskiem jest, iż bardzo niewielka grupa chromosomów po operacji krzyżowania i mutacji oscyluje wokół aktualnego minimum, natomiast bardzo duża grupa osobników nie spełnia żadnych warunków pozwalających na pozytywną ich ocenę. Celem zobrazowania tego rozkładu poniżej przedstawiono wykres zawierający analizę uśrednionych wyników z 50 badań nad zbiorem Wisconsin Breast Cancer po 10 generacjach.

Odchylenie od aktualnego minimum oznacza porównanie jakości osobnika liczonej jako liczba błędnie zaklasyfikowanych przypadków.

Idea przyspieszenia procesu uczenia polegała w tym przypadku na odpowiednio wczesnym „zauważeniu” osobnika nie rokującego nadziei i na zaprzestaniu dalszego sprawdzenia na zbiorze przypadków. Na przykład jeśli aktualny minimalny błąd dopasowania wynosi 30 na 683 przypadki, to osobnik, który po pierwszych

100 przypadkach zakwalifikował błędnie ponad 70, może być uznany za nadający się do odrzucenia już teraz. Oszczędność czasu jest bardzo znaczna! Pytanie, czy tak elitarna strategia selekcji da efekty i jakie parametry wcześniejszej eliminacji będą skuteczne.



Rys. 1. Typowy rozkład jakości osobników w populacji po 10 generacjach

Źródło: opracowanie własne.

Przeprowadzono wiele symulacji, głównie z wykorzystaniem zbioru Wisconsin Breast Cancer. Zastosowano następujący algorytm eliminacji w trakcie badania trafności wnioskowania posiadanych osobników.

Jeżeli minęły już dwie pełne generacje, to:

Jeżeli **liczba sprawdzonych przypadków** przekroczyła 10% **liczby wszystkich przypadków** to:

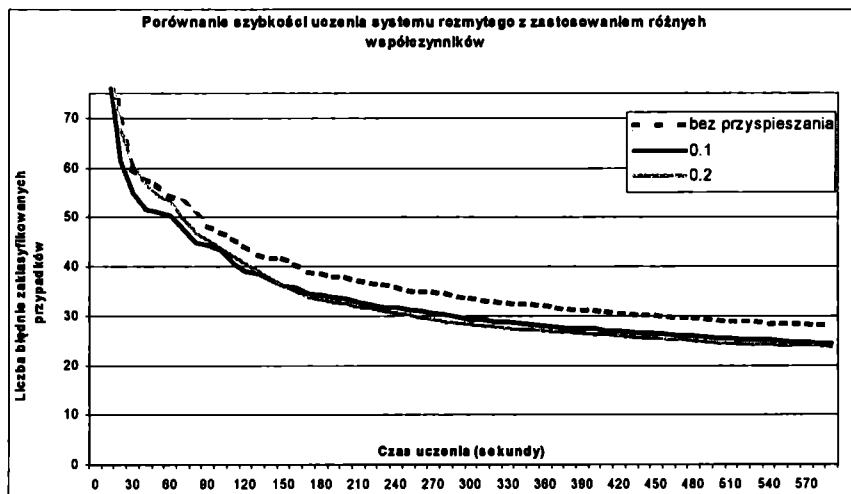
Jeżeli (**aktualny błąd osobnika * współczynnik**) jest większy od (**aktualnego minimum * liczba sprawdzonych przypadków / liczba wszystkich przypadków**) to:

Przerwij proces sprawdzania i przypisz do aktualnego osobnika wysoką wartość błędu.

Przez dwie pierwsze generacje nie stosowano wcześniejszej eliminacji osobników w celu naturalnego ustabilizowania losowo utworzonej populacji za pomocą zwykłej selekcji metodą ruletki. Zmienna **współczynnik** określa próg eliminacji. Im większy, tym szybciej i bardziej zdecydowanie przerywano sprawdzanie osobnika. Oczywiście jest, że ewentualna skuteczność zastosowanego podejścia zależy w dużym stopniu od zastosowanego współczynnika (prog) eliminacji osobników.

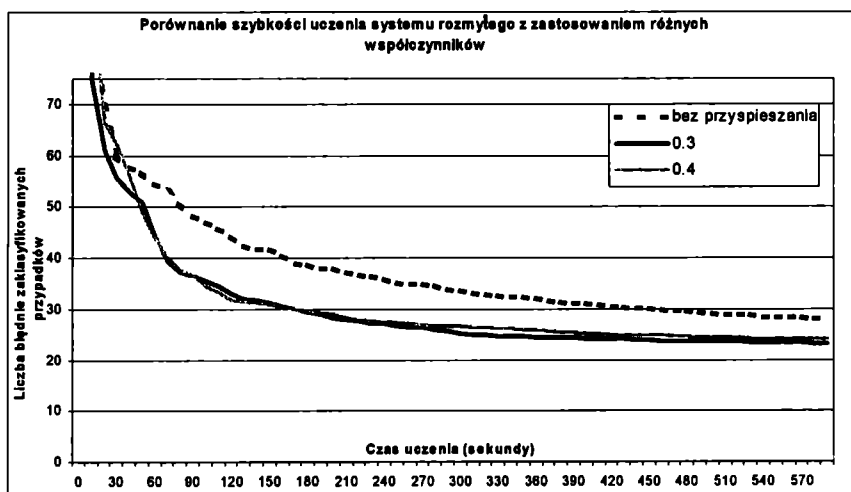
Ponieważ wstępne testy wypadły obiecująco, przeprowadzono cykl badań na zbiorze Wisconsin Breast Cancer. Badania obejmowały po 50 cykli uczących,

trwających po 10 minut¹, dla każdego współczynnika eliminacji. Przebiegi działania AG, przedstawione na rysunkach 2-5, zawierają dane uśrednione.



Rys. 2. Przebieg procesu uczenia na zbiorze WBC ze współczynnikami przyspieszenia 0,1 oraz 0,2

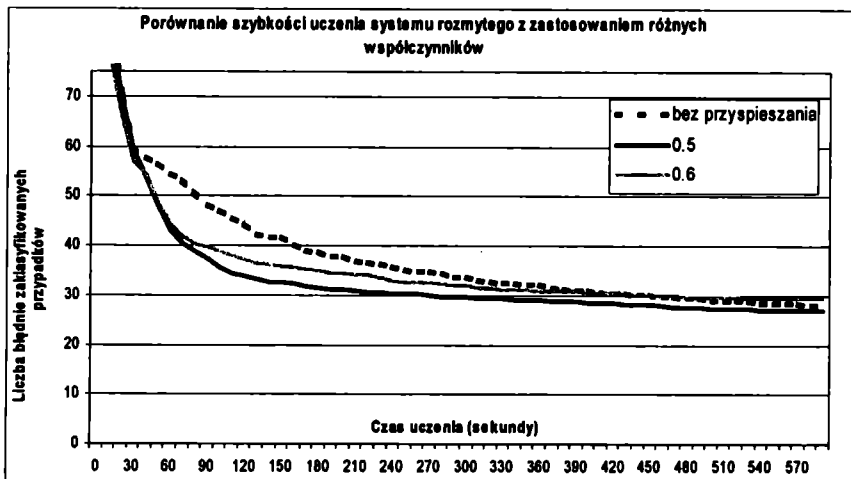
Źródło: badania własne.



Rys. 3. Przebieg procesu uczenia na zbiorze WBC ze współczynnikami przyspieszenia 0,3 oraz 0,4

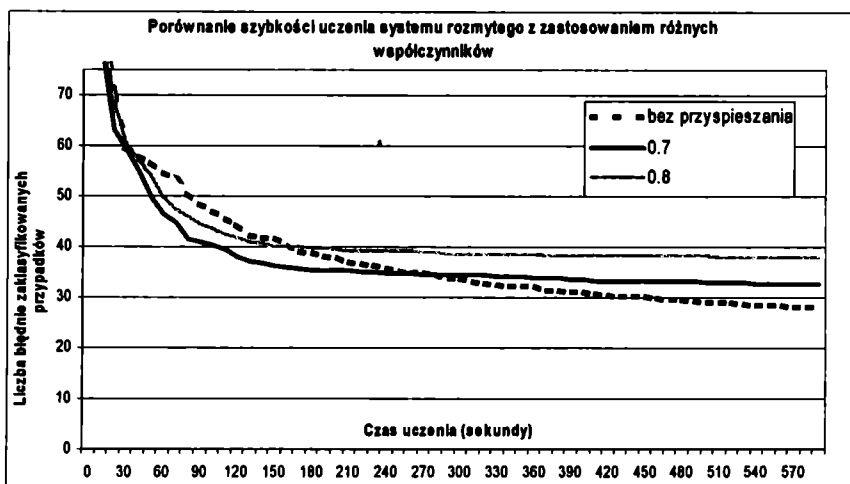
Źródło: badania własne.

¹ Komputer PC z procesorem AMD Duron 1,2 GHz.



Rys. 4. Przebieg procesu uczenia na zbiorze WBC ze współczynnikami przyspieszenia 0,5 oraz 0,6

Źródło: badania własne.

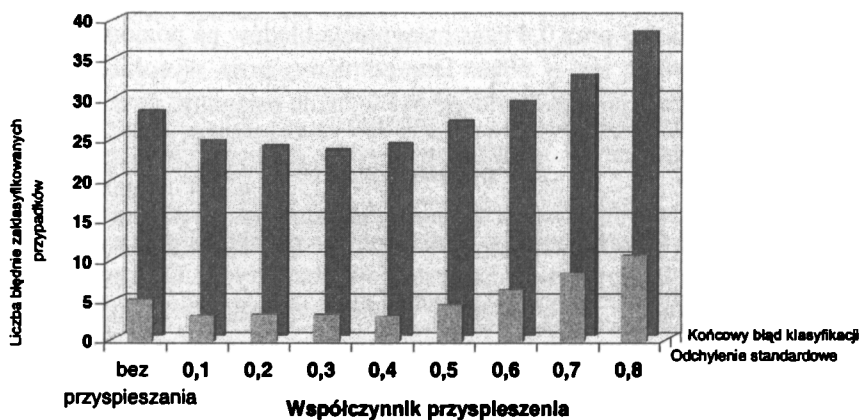


Rys. 5. Przebieg procesu uczenia na zbiorze WBC ze współczynnikami przyspieszenia 0,7 oraz 0,8

Źródło: badania własne.

Analiza wyników badań pokazuje widoczne przyspieszenie pracy AG przy wartościach współczynnika przyspieszenia między 0,1 a 0,6. Zbyt agresywne przyspieszanie, obejmujące współczynniki 0,6 i wyższe, powoduje jednak wczesne zaprzestanie poprawy jakości osobników, co owocuje ostatecznie gorszymi końcowymi wynikami niż przy zastosowaniu konwencjonalnego AG.

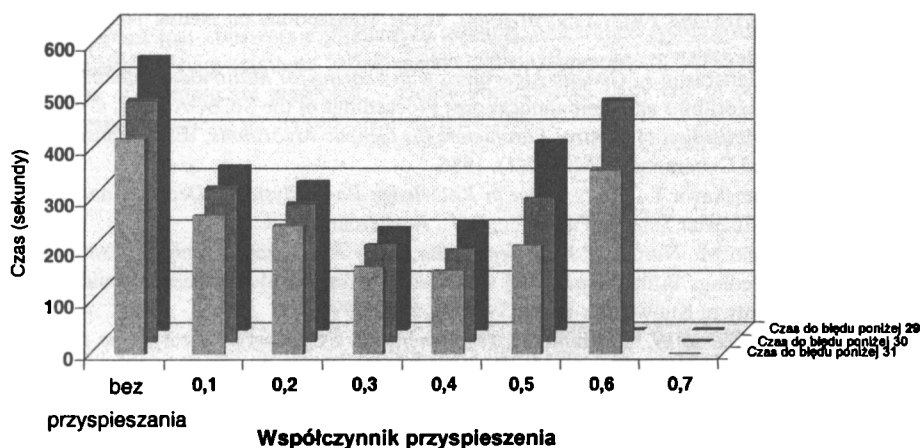
Na rysunku 6 przedstawiono uśrednione końcowe wyniki optymalizowanych systemów.



Rys. 6. Uśrednione końcowe wyniki po 10 minutach uczenia na zbiorze WBC

Źródło: badania własne.

Rysunek 6 zawiera również odchylenie standardowe uzyskanych wyników. Odchylenie standardowe jest istotnym kryterium oceny stabilności pracy AG – im jest mniejsze, tym większym zaufaniem możemy obdarzać uzyskiwane wyniki. Jak widać, oprócz poprawy wyników, istotnie zmalało ich zróżnicowanie w ramach cykli optymalizacyjnych.



Rys. 7. Uśredniony czas osiągnięcia błędów na poziomach: 29, 30 oraz 31, na zbiorze WBC

Źródło: badania własne.

Istotną charakterystykę uzyskanych wyników stanowi okres, jakiego potrzebował AG, aby osiągnąć zakładaną trafność systemu rozmytego. Na rysunku 7 przedstawiono porównanie pracy AG przy różnych współczynnikach przyspieszenia. Dla współczynników 0,3 oraz 0,4 czas osiągnięcia błędów na poziomie 29, 30 oraz 31 przypadków skrócił się o 60%. Dla odmiany, przy wysokich wartościach współczynnika błąd na tym poziomie nie był w ogóle osiągnięty.

7. Podsumowanie

W artykule przedstawiono kilka prób poprawy szybkości pracy algorytmu genetycznego w trakcie optymalizacji systemów rozmytych. Opisywane eksperymenty i modyfikacje nie mają jednakże ścisłego związku z zastosowaniem AG akurat do systemów rozmytych i, w opinii autora, mogą mieć podobny skutek dla innych zastosowań algorytmu, np. do optymalizacji sieci neuronowych czy ekstrakcji reguł. Spośród przedstawionych modyfikacji tylko jedna, polegająca na wcześniejszej eliminacji osobników w populacji, okazała się skuteczna, przynosząc ponad dwukrotne przyspieszenie pracy algorytmu genetycznego.

Obecnie prowadzone są badania obejmujące trudniejsze zbiory danych oraz dalsze modyfikacje AG, polegające m.in. na losowym ustawieniu przypadków w zbiorze dla każdej generacji oraz zmiennego współczynnika przyspieszenia.

Literatura

- [1] Duch W, Adamczak R, Grąbczewski, *A New Methodology of Extraction, Optimization and Application of Crisp and Fuzzy Logical Rules*, IEEE Transactions on Neural Networks nr 12 2001.
- [2] Goldberg D.E. Richardson J., *Genetic Algorithms with Sharing for Multimodal Function Optimization*, Genetic algorithms and their applications: Proceedings of the SICoGA 1987, s. 41-49.
- [3] Grafenstette, *Optimisation of Control Parameters for Genetic Algorithms*, IEEE Transactions on Systems, Man and Cybernetics, SMC-16(1), 1986.
- [4] Gupta M.M., Yamakawa T., *Fuzzy Logic in Knowledge-Based Systems, Decision and Control*, North-Holland / Elsevier Science Publishers B.V., Amsterdam 1988.
- [5] Herrera F., Lozano M., Verdegay J.L., *Generating Fuzzy Rules from Examples Using Genetic Algorithm*, Proceedings Fifth International Conference of Information Processing and Management of Uncertainty in Knowledge-Based Systems, Paris 1994.
- [6] Jin Y., Member, Seelen W., Sendhoff B., *On Generating FC3 Fuzzy Rule Systems from Data Using Evolution Strategies*, IEEE Transactions On Systems, Man, And Cybernetics-Part B: Cybernetics, vol. 29, nr 6, grudzień 1999.
- [7] Kacprzyk J., *Zbiory rozmyte w analizie systemowej*, PWN, Warszawa 1986.
- [8] Michalewicz Z., *Algorytmy genetyczne + struktury danych = programy ewolucyjne*, WNT, Warszawa 1998.
- [9] Nauck D., Nauck U., Kruse R., *Generating Classification Rules with Neuro-Fuzzy System NEFCLASS, NAFIPS'96, Berkeley 1996.*

- [10] Piegat A., *Modelowanie i sterowanie rozmyte*, Akademicka Oficyna Wydawnicza EXIT, Warszawa 1999.
- [11] Schaffer J.D., *A Study of Control Parameters Affecting On-line Performance of Genetic Algorithms for Function Optimization*, 1989.
- [12] Spanos M., Dounias G., Matsatsinis N., Zopounidis C., *A Fuzzy Knowledge Based Decision Aiding Method for the Assessment of Financial Risks: The case of Corporate Bankruptcy Prediction*, European Symposium on Intelligent Techniques, ESIT'99, Chania, Greece, June 1999, (Proc. on CD-Rom, published by the ERUDIT Service Center, Elite Foundation, Germany).
- [13] Wang H.F., Tsaur R.C., *Resolution of Fuzzy Regression Model*, European Journal of Operational Research 2000, vol. 126, nr 1.
- [14] Wasiak M., *Rozmyty system wnioskujący, oceniający kondycję finansową przedsiębiorstw*, [w:] *Pozyskiwanie wiedzy z baz danych*. Prace Naukowe Akademii Ekonomicznej nr 931, AE, Wrocław 2002.
- [15] Wright A.H., *Genetic Algorithms for Real Parameter Optimization*, First workshop on the Foundations of Genetic Algorithms and Classifier Systems, Indiana University, Bloomington, 1990.
- [16] Yager R.R., Filev D.P., *Podstawy modelowania rozmytego*, WNT, Warszawa 1995.
- [17] Yin X. Gernay N., *A Fast Genetic Algorithm with Sharing Scheme Using Cluster Analysis Methods in Multimodal Function Optimization*, Proceedings of International Conference on Artificial Neural Networks and Genetic Algorithms, s. 450-457, Innsbruck 1993.
- [18] Zadeh L.A., *Fuzzy Sets*, Information and Control 1965, nr 8.
- [19] Zieliński J.S., *Inteligentne systemy w zarządzaniu. Teoria i praktyka*, PWN, Warszawa 2000.

SPEEDING UP LEARNING OF FUZZY SYSTEMS USING GENETIC ALGORITHM

Summary

One of the biggest problems of machine learning is the speed of finding patterns, extracting rules, etc. The paper considers a problem of teaching fuzzy logic systems using genetic algorithm, which has also to select key input variables among their numerous set. Different attempts to speed up the process of learning are presented – mainly unsuccessful. Only one of presented methods gave positive results, speeding up the process more than 50%.