



MODELOWANIE PROCESÓW PRODUKCYJNYCH ZA POMOCĄ NAIWNEGO KLASYFIKATORA BAYESA I SZTUCZNYCH SIECI NEURONOWYCH

MARCIN PERZYK, ROBERT BIERNACKI

MODELING OF PRODUCTION PROCESSES BY NAIVE BAYESIAN CLASSIFIER AND ARTIFICIAL NEURAL NETWORKS

Abstract

Modeling qualities of two types learning systems are compared: naive Bayesian classifier (NBC) and artificial neural networks (ANN), based on prediction errors and relevant importance factors of input signals. Simulated and real industrial data were used. It was found that NBC can be an effective and a better tool in some applications, compared to ANN.

1. WSTĘP

Projektowanie i prowadzenie procesów produkcyjnych, wspomagane jest często poprzez wykorzystanie narzędzi matematycznych. W praktyce mamy często do czynienia procesami typu „czarna skrzynka” – o niezidentyfikowanej lub bardzo złożonej naturze fizycznej. Modelowanie tego typu zagadnień polega na znalezieniu zależności między sygnałami wejściowymi a sygnałami wyjściowymi, czyli wynikowymi na podstawie pewnej liczby zaobserwowanych przypadków (zadanie regresji). W ostatnich latach powszechnie wykorzystywanym do tego celu narzędziem są sztuczne sieci neuronowe (SSN), których zastosowanie obejmuje m.in. takie problemy jak:

- przewidywanie własności wyrobów lub materiałów na podstawie parametrów procesu technologicznego,

- przewidywanie awarii urządzeń na podstawie określonych sygnałów,
- identyfikacja przyczyn powstawania wad w wyrobach,
- projektowanie oparte na doświadczeniach zebranych w przemyśle, uogólnionych przez SSN.

Jednym z głównych problemów dotyczących wykorzystania SSN jest to, że wyniki modelowania nie są jednoznaczne. Postawione zadanie mogą realizować sieci o różnej architekturze, topologii połączeń, a ponadto dla danej sieci poszczególne uczenia prowadzą do różnych wartości wag synaps, związanych z osiąganiem minimów lokalnych błędów sieci (Perzyk i in., 2003; Perzyk i Kochański, 2003).

Taka niejednoznaczność przewidywań nie występuje w przypadku modelowania z wykorzystaniem innych narzędzi matematycznych, np. aproksymacji przy pomocy wielomianów oraz, będącego przedmio-

M. Perzyk, R. Biernacki, Instytut Technologii Materiałowych Politechniki Warszawskiej, M.Perzyk@wip.pw.edu.pl



tem niniejszej pracy, naiwnego klasyfikatora Bayesa (NKB) (Cichosz, 2000) - prostego systemu uczącego się wykorzystującego rachunek prawdopodobieństwa. Autorom nie są znane jego zastosowania do modelowania procesów technologicznych.

Badania przedstawione w pracy dotyczą analizy przydatności NKB do zagadnień modelowania procesów produkcyjnych, obejmującej porównanie błędów przewidywań tego systemu z błędami SSN, a także tzw. współczynników istotności względnych sygnałów wejściowych dla danych modelowanych za pomocą obu systemów.

2. SYSTEMY UCZĄCE SIĘ WYKORZYSTUJĄCE NAIWNY KLASYFIKATOR BAYESA

Systemy wykorzystujące metody probabilistyczne oparte są najczęściej na twierdzeniu i wzorze Bayesa. Należą do nich: klasyfikatory Bayesa (optymalny, bez znaczenia praktycznego oraz NKB), a także sieci Bayesowskie (Cichosz, 2000).

Twierdzenie Bayesa jest stosowane do zdarzeń o znanym wyniku, których zajście było zależne od czynników poprzedzających wystąpienie tych zdarzeń. Obliczane jest prawdopodobieństwo, na ile dany czynnik stanowi przyczynę wystąpienia zdarzenia (skutku). Zastosowanie NKB jest związane z obliczeniami prawdopodobieństw na podstawie odpowiednio przygotowanego zbioru uczącego, składającego się z przykładów opisanych za pomocą atrybutów i pojęcia docelowego. Zastosowanie tego narzędzia sprowadza się do następujących działań:

- Ustalenie atrybutów (kategorii) dla wielkości wejściowych oraz kategorii wielkości wyjściowej (inaczej pojęcia docelowego). Należy zauważyć, że NKB wymaga stosowania wielkości o charakterze nominalnym lub porządkowym, a nie ciągłym. „Wartość” atrybutu lub pojęcia docelowego jest określana przez oznaczenie jego przynależności do danej kategorii, która może być identyfikowana werbalnie, np. coś może być „bardzo małe”, „małe”, „średnie”, „duże” i „bardzo duże” (5 kategorii), albo „dobre” i „złe” (2 kategorie).
- Utworzenie zbioru uczącego, składającego się z zapisów (rekordów) zawierających wartości atrybutów wszystkich wielkości wejściowych oraz odpowiadającej im wartości kategorii wielkości wyjściowej.
- Na podstawie zbioru uczącego szacuje się prawdopodobieństwa poszczególnych kategorii wielkości wyjściowej oraz prawdopodobieństwa poszczególnych wartości atrybutów (kategorii) wszystkich wielkości wejściowych dla poszczególnych kate-

gorii wyjścia. Etap ten jest określany jako uczenie (trenowanie) klasyfikatora.

- Na podstawie prawdopodobieństw oszacowanych na etapie uczenia, można obliczyć prawdopodobieństwa wystąpienia każdej z kategorii wielkości wyjściowej dla dowolnego przypadku (zestawu wielkości wejściowych). Odpowiedzią NKB jest ta wartość kategorii wielkości wyjściowej, która uzyskała największą wartość prawdopodobieństwa. W przypadku, gdy kategorie wyjścia zostały utworzone przez przypisanie liczbom oznaczającym nazwy tych kategorii przedziały wartości wyjścia ciągłego, możliwa jest zamiana nazwy obliczonej przez NKB kategorii na wartość ciągłą, obliczoną np. jako średnią z takiego przedziału.

3. METODYKA BADAŃ

3.1. Zbiory danych

W badaniach zastosowano zbiory danych różniące się charakterem zależności między sygnałami wejścia i wyjścia, zarówno uzyskane w przemyśle, jak i sztucznie wygenerowane. Te drugie, o znanych zależnościach pomiędzy wejściem a wyjściem, pozwalają na lepszą ocenę poprawności przewidywań systemów uczących się. Zbiory danych symulowanych otrzymywano na podstawie założonych prostych wzorów typu $Y = f(X_1, X_2, \dots)$, wprowadzając losowe zakłócenia w zakresie $\pm 20\%$, w sposób opisany szczegółowo w Perzyk i in., 2003). W niniejszej pracy wykorzystano następujące zbiory danych:

- *Dane symulowane 1* – zbiór utworzony sztucznie na podstawie wzoru: $Y = X_1 + 2 \cdot X_2 + 3 \cdot X_3 + 4 \cdot X_4 + 5 \cdot X_5$. Liczebność podzbioru uczącego 1000 rekordów, zaś weryfikującego 200.
- *Dane symulowane 2* – zbiór wygenerowany wg wzoru: $Y = (6 \cdot X_1)^3 + (10^{-3} \cdot X_2)^3 + \dots + (1 \cdot X_7)^3 + \dots + (1 \cdot X_{12})^3$. Utworzone wartości wyjściowe tego zbioru mają charakter ciągły. Podzbiór uczący zawiera 1000 rekordów, weryfikujący 200 rekordów.
- *Dane symulowane 3* – wygenerowane wartości wyjściowe tego zbioru mają charakter binarny 0 lub 1, zależnie od przekroczenia wartości granicznej Y , przy czym wartości danych wejściowych były identyczne jak dla zbioru symulowanego 2. Wartość graniczną dobrano tak, że udział wartości $Y = 1$ w zbiorze wynosił ok. 18%, co odpowiadało udziałowi braków w danych przemysłowych 2 (opisanych poniżej). Liczebność jak zbioru powyższego.
- *Dane symulowane 4* – zbiór uczący utworzony wg wzoru jak dla danych symulowanych 3, jednak o niewielkiej liczebności, takiej jak zbiór danych przemysłowych 2 (172 rekordy, zbioru weryfikującego nie utworzono).

- *Dane przemysłowe 1* – wyniki badań wytrzymałości na rozciąganie żeliwa sferoidalnego w zależności od jego składu chemicznego, określonego przez 9 pierwiastków, uzyskane jako wynik procesu wytopu prowadzonego w piecach indukcyjnych w jednej z odlewni żeliwa (Perzyk i Kochański, 2001). Oprócz podzbioru uczącego liczącego 700 rekordów utworzono podzbiór weryfikujący zawierający 90 rekordów.
- *Dane przemysłowe 2* – wielkościami wejściowymi są parametry procesu produkcyjnego związane z formą piaskową (12 parametrów), zaś wielkością wyjściową – wystąpienie wady typu porowatość gazowa w odlewie stalowym (Perzyk i Kochański, 2001). Wyniki uzyskano w jednej z odlewni krajowych produkującej odlewy w krótkich seriach, w formach z mas glinowych wilgotnych. W tym przypadku oryginalne wyjście było typu nieciągłość (kategoria „1” – brak wady, kategoria „2” – wystąpienie wady). Zbiór uczący zawierał 172 rekordy, zbioru weryfikującego nie było.

Dane symulowane 1 zaplanowano w ten sposób, ażeby porównać m.in. precyzję przewidywania istotności względnych sygnałów wejściowych dla obu systemów. *Dane symulowane 2, 3 i 4* zaprojektowano tak, ażeby umożliwić analizę porównawczą niektórych wyników dla tych zbiorów z wynikami dla *danych przemysłowych 2*. Zbiory te odnoszą się do hipotetycznej sytuacji, w której staramy się odkryć, które z 12 parametrów procesu mogą być przyczyną pogorszenia się jakości wyrobu. *Zbiór symulowany 2* zakłada, że parametrem jakości jest wielkość o charakterze ciągłym – może to być np. chropowatość powierzchni albo jakaś własność materiału. *Zbiory symulowane 3 i 4* naśladują typową sytuację, w której wyroby klasyfikowane są jako dobre lub wadliwe. We wszystkich trzech założono, że decydujący jest sygnał X1, ale również suma działań 6 sygnałów (od X7 do X12) może dać ten sam efekt; sygnały od X2 do X6 są natomiast praktycznie bez znaczenia. Z założeń tych wynikają wartości stałych współczynników we wzorze użytym do generacji, natomiast zastosowanie 3-cich potęg miało na celu uzyskanie silnej nieliniowości, charakterystycznej dla procesów produkcyjnych.

Utworzone zbiory uczące były wykorzystywane we wcześniejszych badaniach do modelowania określonych zależności z wykorzystaniem sztucznych sieci neuronowych.

3.1. Sieci neuronowe

Szczegóły dotyczące sieci neuronowych, w tym ich topologii, metod uczenia oraz sposobu wykorzystania zbiorów weryfikujących, znaleźć można w pracy

(Perzyk i in., 2003). Wykorzystywana w niniejszej pracy definicja współczynników istotności względnej sieci oparta jest na wielkości wzrostu błędu sieci dla danych uczących wskutek zablokowania danego wejścia na stałym poziomie. Dla wszystkich rekordów zbioru uczącego wyliczane są i sumowane różnice pomiędzy modułami błędów względnych (tj. odniesionych do wartości rzeczywistych) a modułami błędów przewidywania sieci dla wartości danego wejścia równej minimalnej. Otrzymane w ten sposób sumy dla wszystkich wejść, są normalizowane w zakresie 0 - 1 tak, że najbardziej znacząca wielkość ma przypisaną wartość 1, zaś najmniej znacząca - 0 i stanowią wartości współczynników istotności wielkości wejściowych sieci. Dyskusję na temat znaczenia tych współczynników oraz ich ograniczeń znaleźć można w pracy (Perzyk i in., 2003).

3.2. Klasyfikator Bayesa

W niniejszej pracy zmienne ciągle zamieniano na wartości atrybutów i kategorie przez zaliczenie danej zmiennej do oznaczonego numerem odpowiedniego przedziału tej oryginalnej wartości. Dla każdego ze zbiorów danych wykonano próby z dobieraniem liczby przedziałów (kategorii) dla wielkości wejściowych oraz wielkości wyjściowych o wartościach typu ciągłego, przy czym w danym uczeniu stosowano dla wszystkich zmiennych jednakową liczbę przedziałów. Warto zauważyć, że mała liczba kategorii zmniejsza precyzję klasyfikatora, zarówno przy uczeniu się, jak odpytywaniu. Duża liczba kategorii stwarza z kolei niebezpieczeństwo, że niektóre z nich będą nie reprezentowane w zbiorze uczącym, a przynajmniej reprezentacje będą bardzo mało liczne, nawet równe zero. Wprawdzie sposób szacowania prawdopodobieństw warunkowych dla klasyfikatora uwzględnia takie sytuacje (m.in. automatyczne tworzenie „wirtualnych” przypadków (Cichosz, 2000)), jednakże jakość takiego uczenia się będzie z pewnością gorsza.

Warto podkreślić, że autorom nie są znane jakiegokolwiek wskazówki dotyczące dyskretyzacji wielkości ciągłych na potrzeby klasyfikatora Bayesa. W przypadku podziału zbioru wartości ciągłych na przedziały w celu budowy histogramów uważa się, że liczba przedziałów powinna zawierać się w granicach od $3/4 \sqrt{n}$ do \sqrt{n} , gdzie n oznacza liczebność zbioru. W niniejszej pracy celowo przyjęto szersze granice liczb przedziałów.

Wszystkie obliczenia związane z NKB zaprogramowano w języku VBA for Excel i przedstawiano w arkuszu kalkulacyjnym. Warto zwrócić uwagę, że zbiory weryfikujące nie są potrzebne do uczenia NKB, a jedynie mogą służyć do oceny jego zdolności przewidywania dla nowych danych, jako zbiory testujące.

Na rys. 1 pokazano wyniki prób podziału danych na kategorie. Pokazano wykresy względnych błędów średnich wyników modelowania dla zbiorów o różnej charakterystyce. W niniejszej pracy dla zbiorów z wyjściem ciągłym błędy te definiowano jako:

$$\frac{|(\text{odpowiedź systemu} - \text{wartość rzeczywista})|}{\text{wartość rzeczywista}} \times 100\%$$

W przypadku danych z wyjściem wyrażonym w kategoriach, jak dla *danych przemysłowych 2* (a także *danych symulowanych 3*, nie przedstawionych na tym rysunku) błąd definiowano jako procent mylnych kategorii, tzn. stosunek liczby przypadków w danym zbiorze, dla których system przewiduje kategorię różną od rzeczywistej, do liczebności zbioru.

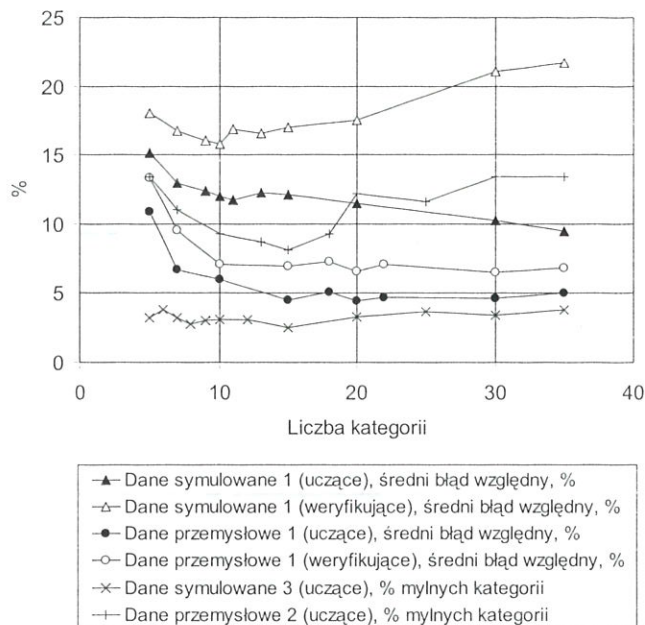
Rozkłady błędów dla danych uczących i weryfikujących są zbliżone, generalnie najniższe błędy uzyskuje się dla ilości kategorii tym mniejszych im mniej liczny jest zbiór uczący. Przyjęto liczby kategorii dające najmniejsze błędy dla danych weryfikujących (lub uczących w przypadku braku weryfikujących), zapewniające jednak minimalną liczbę wystąpień (dowolnej kategorii i dowolnego wejścia lub wyjścia) równą 2.

Należy podkreślić, że przyjęta w pracy, w znacznym stopniu intuicyjnie, metodyka jest tylko jedną z możliwych i dalsze prace nad nią są z pewnością wskazane.

Współczynniki istotności względnej dla NKB obliczono na podstawie analizy danych uczących wyrażonych w kategoriach, dla liczb kategorii wybranych jako optymalne dla uczenia klasyfikatora w sposób opisany powyżej. Jako miarę istotności danego wejścia przyjęto średni przyrost wyjścia obliczany w następujący sposób: ze zbioru uczącego wybierane są te zapisy, w których dane wejście przyjmuje jedną, stałą wartość np. kategorię o nazwie „1”, a następnie wyliczane są różnice wyjść pomiędzy tymi rekordami, a pozostałymi (tj. takimi, w których dane wejście ma inną wartość), na zasadzie wszystkich możliwych kombinacji par rekordów. Następnie wybierana jest kolejna ustalona wartość (kategoria) danego wejścia i obliczana jest kolejna seria takich różnic, a w końcu wyznacza się z nich średnią różnicę dla danego wejścia. Takie postępowanie powtarza się dla wszystkich wejść, zaś ostateczne wartości współczynników istotności względnych wyznaczane są przez znormalizowanie tych średnich wg zasady: wartość minimalna wynosi 0, zaś maksymalna równa jest 1. Należy zaznaczyć, że zastosowano tu metodykę obliczeń bez bezpośredniego wykorzystania przewidywań klasyfikatora; podejście analogiczne jak dla współczynników istotności sieci neuronowych nie dało pozytywnych wyników.

4. PORÓWNANIE BŁĘDÓW NKB I SSN

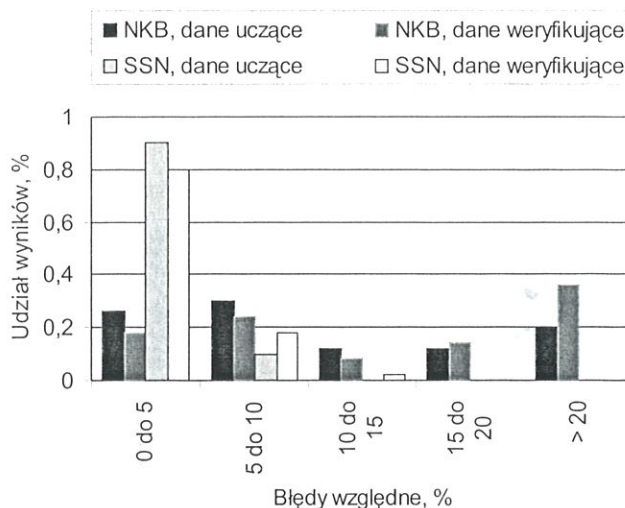
Określono błędy względne wyników przewidywa-



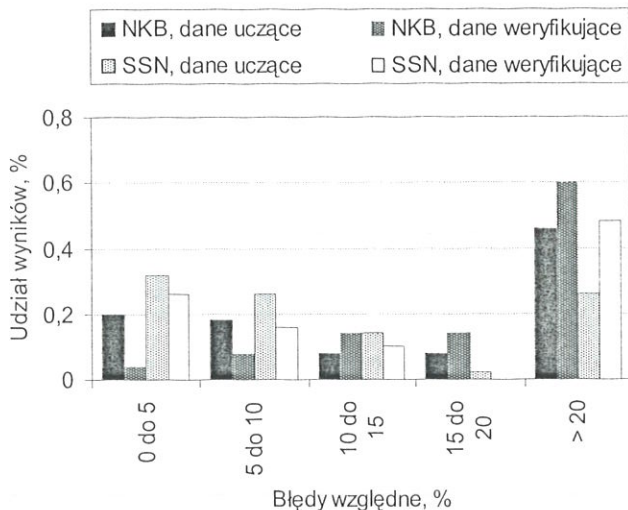
Rys. 1. Błędy przewidywania wielkości wyjściowej przez NKB dla wybranych zbiorów danych, w funkcji liczby kategorii

nia wielkości wyjściowych z wykorzystaniem NKB i SSN, uwzględniając ich udział w charakterystycznych przedziałach wartości. Na rys. 2 pokazano rozkłady błędów przewidywania dla *danych symulowanych 1*. Z rysunku wynika, że SSN zdecydowanie dokładniej modeluje zależności opisane tego typu zbiorem danych, tj. o „wyraźnych” zależnościach i stosunkowo dużej liczebności. Stosunkowo duże wartości błędów obu systemów, w tym SSN, są niewątpliwie wynikiem wprowadzonego celowo znacznego szumu w danych uzyskiwanych z symulacji.

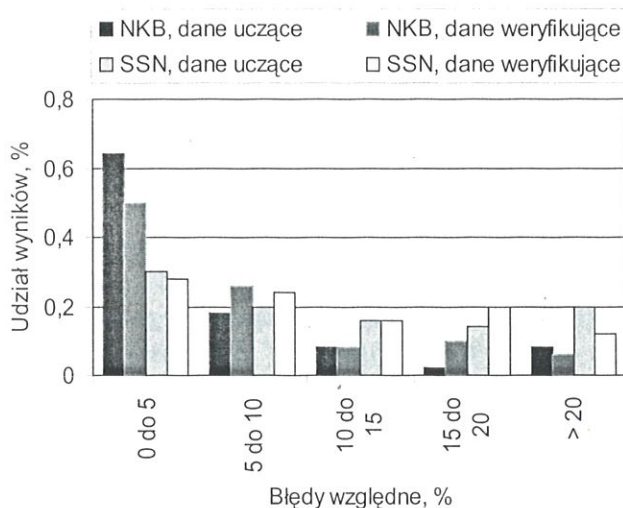
Rys. 3 przedstawia rozkład błędów dla zbioru symulowanego naśladującego w pewnym stopniu charakter *danych przemysłowych 2*, ale o wyjściu typu



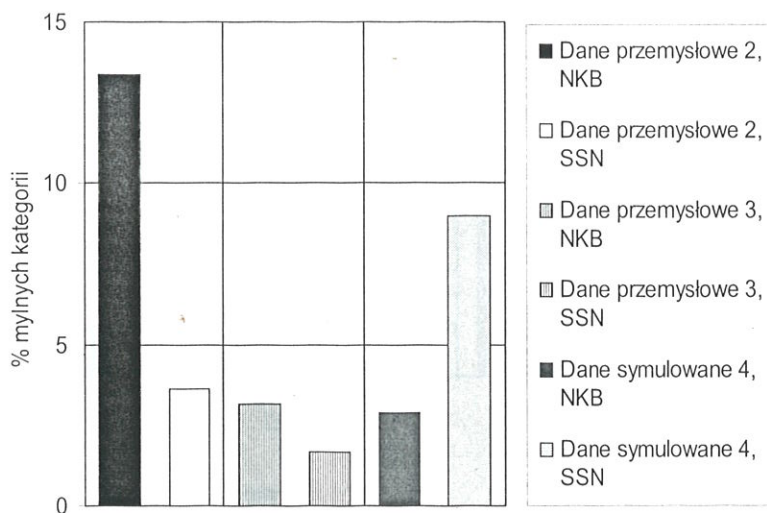
Rys. 2. Rozkłady błędów przewidywania wielkości wyjściowej przez NKB i SSN (minimalne błędy z 10 uczeń) dla danych symulowanych 1



Rys. 3. Rozkłady błędów przewidywania wielkości wyjściowej przez NKB i SSN (minimalne błędy z 10 uczeń) dla danych symulowanych 2



Rys. 4. Rozkłady błędów przewidywania wielkości wyjściowej przez NKB i SSN (minimalne błędy z 10 uczeń) dla danych przemysłowych 1



Rys. 5. Rozkłady błędów przewidywania wielkości wyjściowej przez NKB i SSN (średnie błędy z 10 uczeń) dla danych z wyjściem binarnym

ciągłego. Przewidywania obu systemów modelujących założone relacje są obarczone wysokimi błędami, należy jednak podkreślić, że SSN nieco lepiej poradziła sobie z wykonywanym zadaniem, gdyż udział błędów małych jest dla tego systemu większy.

W przypadku *danych przemysłowych 1* – o wyjściu ciągłym (uzyskane z odlewni wyniki badań wytrzymałości żeliwa sferoidalnego w zależności od jego składu chemicznego, określonego przez 9 pierwiastków) NKB wyraźnie lepiej przewiduje wartości wyjściowe niż SSN (rys.4). Świadczy to, że większa precyzja przewidywań SSN nie musi być regułą.

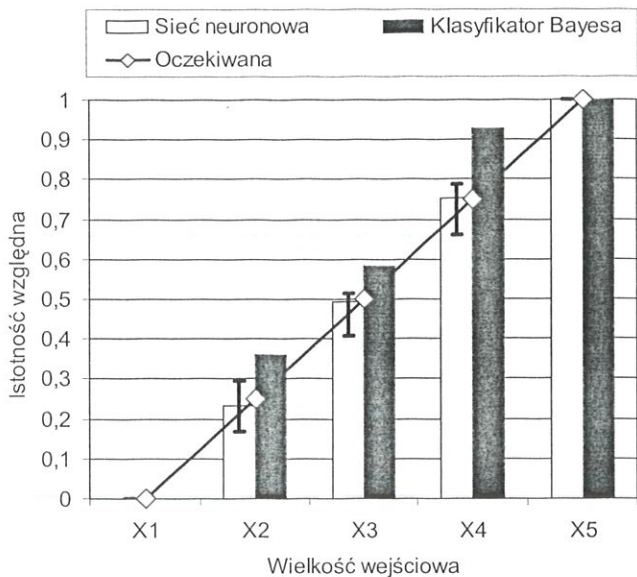
Rys. 5 przedstawia porównanie rozkładu błędów wyrażonych w procentach mylnych kategorii dla zbiorów uczących z wyjściem binarnym. Ciągłe odpowiedzi SSN przeliczono do postaci binarnej w ten sposób, że wartościom z przedziału 0–0,5 przypisano oznaczenie – „kategorię” 1, zaś wartościom z przedziału 0,5–1 przypisano oznaczenie 2. Widoczne jest, że SSN zdecydowanie lepiej niż NKB modeluje zależności opisane wyjściem typu binarnego dla danych przemysłowych 2, zaś dla danych symulowanych zależnie od liczebności zbioru uczącego: NKB charakteryzuje się wynikami lepszymi dla zbiorów małych.

5. WSPÓŁCZYNNIKI ISTOTNOŚCI WZGLĘDNEJ SYGNAŁÓW WEJŚCIOWYCH

Jedną z korzyści, jakie może dać modelowanie procesów produkcyjnych jest wskazanie wielkości odgrywających dominującą czy przynajmniej bardzo znaczącą, albo też znikomą bądź żadną, rolę w procesie. W pierwszym przypadku jest to pomocne np. w wykrywaniu przyczyn awarii urządzeń, zakłóceń procesów technologicznych, określaniu parametrów procesu technologicznego decydujących o jakości wyrobu i w innych tego typu sytuacjach. W drugim może służyć do redukcji liczby wielkości wejściowych, co m. in. oznacza uproszczenie procedur rejestrowania i kontrolowania parametrów procesu.

Na rys. 6–9 pokazano wykresy istotności sygnałów wejściowych, oszacowanych za pomocą metod omówionych w rozdziale 3. Dla SSN pokazano także rozrzuty wyników z 10 uczeń, zdefiniowane jako zakresy pomiędzy minimalną a maksymalną wartością.

W przypadku pokazanym na rys. 6 wyliczone istotności względne generalnie zgodne są z oczekiwaniami (rozkład istotności względnych odzwierciedla narastające znaczenie kolejnych wielkości wejściowych od 1 do 5-tej). Jednakże SSN jest w stanie prze-



Rys. 6. Istotności względne wielkości wejściowych dla NKB, SSN oraz oczekiwane dla danych symulowanych 1 (dla SSN pokazano za pomocą pionowych linii także rozrzuty wyników z 10 uczeń)

widywać znaczenie sygnałów wejściowych o wiele precyzyjniej niż NKB.

Dla danych przemysłowych 1 (wyników badań dotyczących wytrzymałości żeliwa sferoidalnego, rys. 7), NKB prawidłowo wyróżnił miedź, która jest pierwiastkiem stosowanym w praktyce produkcyjnej do zwiększenia wytrzymałość żeliwa sferoidalnego (SSN wskazała tę wielkość jako drugą pod względem znaczenia). Zwraca uwagę to, że wartości współczynników istotności względnych z wykorzystaniem wyników modelowania SSN charakteryzują się znacznym rozrzutem.

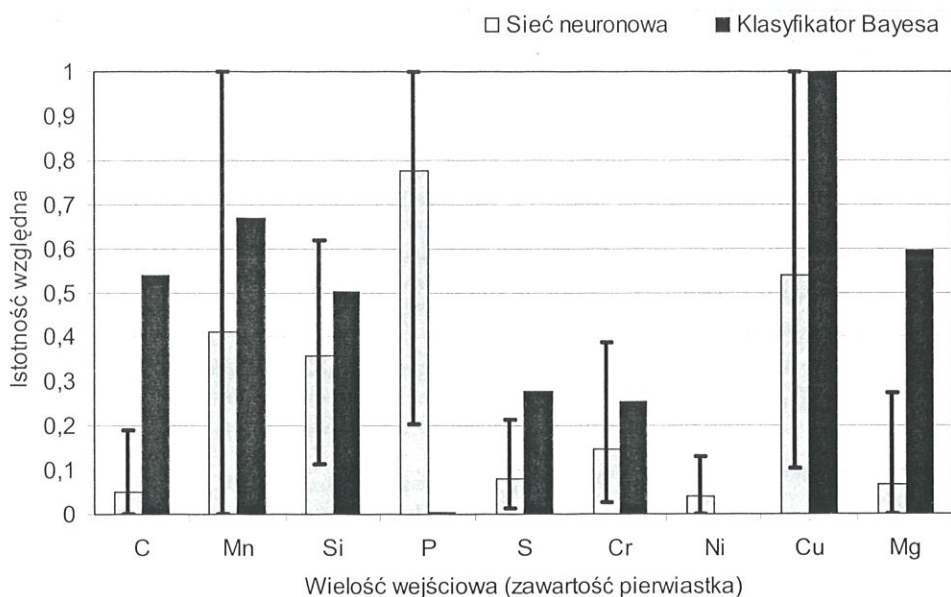
Rysunki 8 i 9 dotyczą rozkładów istotności zbiorów symulowanych o takiej samej strukturze danych

wejściowych, różniących się typem danych wyjściowych (rys. 8 – dane symulowane 2, o wyjściu ciągłym; rys. 9 – dane symulowane 3, o wyjściu binarnym). Jak wspomniano, oba zbiory danych odnoszą się do hipotetycznej sytuacji, w której staramy się odkryć, które z 12 parametrów procesu mogą być przyczyną pogorszenia się jakości wyrobu. Pierwszy z nich zakłada, że parametrem jest wielkość o charakterze ciągłym – może to być np. chropowatość powierzchni albo jakaś własność materiału. Drugi zbiór naśladuje typową sytuację, w której wyroby klasyfikowane są jako dobre lub wadliwe. W obu zestawach danych założono, że decydujący jest sygnał X1, ale również suma działań 6 sygnałów (od X7 do X12) może dać ten sam efekt; sygnały od X2 do X6 są natomiast praktycznie bez znaczenia.

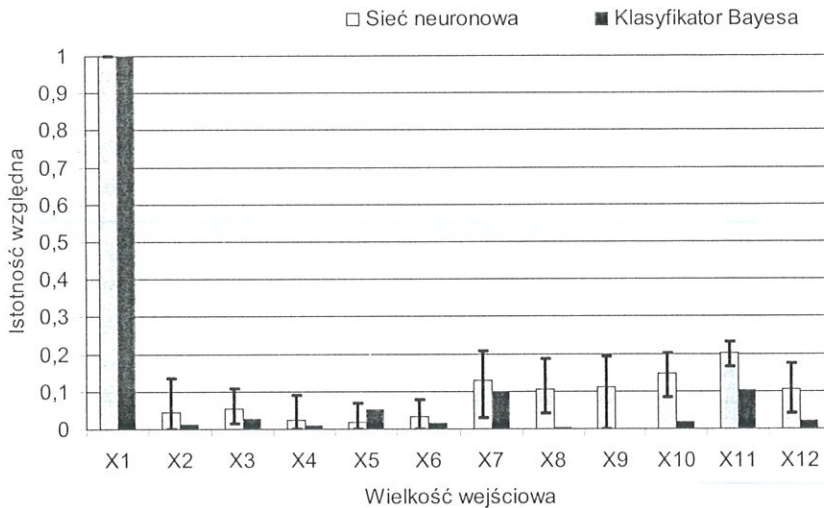
Otrzymane współczynniki istotności względnej dla zbioru z wyjściem ciągłym (rys. 8) potwierdzają fakt zdecydowanie wyróżniającego się działania jednego sygnału X1 zarówno przez SSN i NKB, a w przypadku SSN także generalnie nieco większe istotności sygnałów X7 ... X12 w stosunku do sygnałów X2 ... X6.

Dla wyjść o charakterze binarnym (dane symulowane 3, rys. 9) widoczne jest dramatyczne zwiększenie rozrzutów wyników dla poszczególnych uczeń SSN w porównaniu z rys. 8, do rozmiarów uniemożliwiających praktyczne wykorzystanie współczynników istotności. Wyniki klasyfikacji związanej z NKB mają natomiast charakter jednoznaczny. Warto podkreślić, że niejednoznaczność przewidywań SSN, przejawiająca się m.in. w dużych rozrzutach istotności względnych wielkości wejściowych, wydaje się być nie do uniknięcia, pomimo że poczyniono wiele wysiłków, aby ją ograniczyć (patrz np. Yescas, 2001).

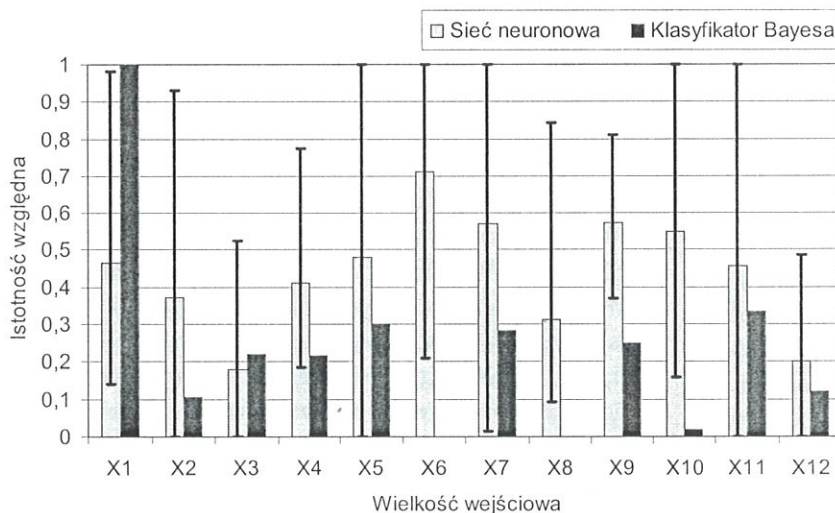
Należy zaznaczyć, że obecna postać współczynników istotności wielkości wejściowych nie ma sensu fizycznego, a jedynie pozwala wskazanie sygnałów najważniejszych, mogących być np. potencjalnymi przyczynami zakłóceń procesu technologicznego, jak również najmniej istotnych, których pominięcie można by rozważyć przy kształtowaniu systemu uczącego się. Podobnie jak w innych pracach (np. Yescas, 2001), współczynniki istotności nie uwzględniają współzależności i efektów synergii wielkości wejściowych.



Rys. 7. Istotności względne wielkości wejściowych dla zbioru danych przemysłowych 1



Rys. 8. Istotności względne wielkości wejściowych dla zbioru danych symulowanych 2



Rys. 9. Istotności względne wielkości wejściowych dla zbioru danych symulowanych 3

Pewne prace w tym kierunku, w odniesieniu do SSN, przedstawiono w (Perzyk i in., 2004).

6. PODSUMOWANIE I WNIOSKI

Badania porównawcze przeprowadzone w niniejszej pracy pozwoliły na stwierdzenie, że błędy przewidywania naiwnego klasyfikatora Bayesa być mniejsze lub większe od błędów sieci neuronowej. Dla dwóch identycznych symulowanych zbiorów o wyjściu binarnym, różniących się tylko liczebnością, NKB wykazał mniejszy błąd (% mylnych kategorii) niż SSN w przypadku zbioru mniej liczniejszego.

Współczynnik istotności względnej wielkości wej-

ściowych związany z NKB w obecnej postaci daje mało precyzyjne wyniki dla sygnałów o mniejszym znaczeniu. Dla wyjścia typu binarnego NKB dał wyraźnie lepszy rezultat w postaci identyfikacji sygnału zdecydowanie wyróżniającego się, niż SSN.

Należy stwierdzić, że główną zaletą NKB w stosunku do SSN jest jednoznaczność tego modelu, a także prostota jego stosowania. Sądzić można także, że NKB jest mniej wymagający, jeśli chodzi o liczebność zbioru uczącego (np. nie wymaga on zbioru weryfikującego przy uczeniu).

Wydaje się, że NKB może stanowić pożyteczny dla zastosowań przemysłowych system uczący się, w niektórych przypadkach lepszy od SSN. Celowe są jednak dalsze, systematyczne badania, zwłaszcza nad rozszerzeniem możliwości interpretacji wyników obu systemów, w tym analizy istotności wielkości wejściowych. Obecna praca pozwoliła jedynie na zwrócenie uwagi na potencjalne możliwości zastosowań NKB do modelowania procesów przemysłowych, jako alternatywnego wobec SSN systemu uczącego się oraz zasygnalizowanie szeregu problemów z tym związanych.

LITERATURA

- Cichosz, P.: *Systemy uczące się*. WNT Warszawa 2000.
- Perzyk, M., Kochański, A., 2001, Prediction of ductile cast iron quality by artificial neural networks. *J. Mat. Proc. Techn.*, 109, 305-307.
- Perzyk, M., Kochański, A., Kozłowski, J., 2003, Istotność względna sygnałów wejściowych sieci neuronowej. *Informatyka w Technologii Materiałów*, 3, 125 - 132.
- Perzyk, M., Kochański, A., 2003, Detection of causes of casting defects assisted by artificial neural networks. *J. Eng. Manufacture, Proc. of the Institution of Mechanical Engineers*, Part B, 217, 1279 - 1284.
- Yescas, M.A., Bhadeshia, H.K.D.H., MacKay, D.J., 2001, Estimation of the amount of retained austenite in austempered ductile irons using neural networks, *Mat. Sci. Eng.*, A311, 162 ? 173.
- Perzyk, M., Kochański, A., Kozłowski, J., 2004, Significance analysis of production parameters assisted by artificial neural networks. *Proc. Int. Conf. Advances in Production Engineering*, Warsaw University of Technology, Warsaw.

Artykuł otrzymano 4 czerwca 2004 r.