

SIECI NEURONOWE JAKO MODELE PROGNOZOWANIA UŻYCIA NARZĘDZI W SYSTEMACH CAPP

IZABELA ROJEK

Uniwersytet Kazimierza Wielkiego w Bydgoszczy

Streszczenie

W referacie przedstawiono sieci neuronowe jako modele prognozowania użycia narzędzi w systemach CAPP. Modele opracowano przy użyciu liniowej sieci neuronowej, sieci jednokierunkowej wielowarstwowej ze wsteczną propagacją błędów oraz sieci o radialnych funkcjach bazowych. Dokonano porównania tych modeli. Miarą oceny modeli jest skuteczność przewidywania użycia narzędzi w różnych interwałach czasowych. Modele te są wykorzystane już na etapie projektowania procesu technologicznego, w celu planowego wytworzenia wyrobów i aby nie doszło do przestoju z powodu braku narzędzi, szczególnie narzędzi specjalnych.

Utworzone modele przetestowano na danych rzeczywistych z przedsiębiorstwa produkcyjnego.

Słowa kluczowe: sieć neuronowa, model predykcji, narzędzie

1. Wstęp

Współcześnie systemy komputerowego wspomaganie projektowania procesów technologicznych (systemy CAPP) coraz częściej zawierają metody sztucznej inteligencji wspomagające projektowanie procesu technologicznego.

W trakcie projektowania procesu technologicznego bardzo ważnym etapem jest dobór narzędzi, ze względu na dużą różnorodność możliwości doboru w trakcie samego projektowania, jak i w sytuacjach zmiany narzędzi w trakcie procesu wytwarzania. Na tym etapie staje się niezbędne doświadczenie technologa.

W celu polepszenia doboru narzędzi do operacji technologicznej opracowano:

- bazę danych, wykorzystywaną do opracowania modeli i metod inteligentnego wspomaganie,
- model doboru narzędzi do operacji technologicznych,
- model przewidywania użycia narzędzi w różnych interwałach czasowych,
- model preferencji technologa w doborze narzędzi.

W celu zapewnienia zrealizowania planu produkcji niezbędne w systemach CAPP jest opracowanie modeli przewidywania użycia narzędzi w różnych interwałach czasowych. Te modele są wykorzystane już na etapie projektowania procesu technologicznego, aby w odpowiednim czasie zostały wytworzone wyroby i nie doszło do przestoju z powodu braku narzędzi, szczególnie narzędzi specjalnych.

Miarą oceny modeli jest skuteczność przewidywania użycia narzędzi w różnych interwałach czasowych. Utworzone modele przetestowano na danych rzeczywistych z przedsiębiorstwa produkcyjnego. Opracowane modele mogą być szczególnie przydatne i ważne dla przedsiębiorstw charakteryzujących się produkcją jednostkową, o bardzo małej powtarzalności oraz bardzo dużej precyzji wykonania.

2. Sieć neuronowa jako model prognostyczny

Sieci neuronowe zwracają na wyjściu wartości ciągłe, dlatego znakomicie nadają się do szacowania i przewidywania. Sieci takie mogą analizować jednocześnie wiele zmiennych. Możliwe jest stworzenie modelu, nawet gdy rozwiązanie ma dużą złożoność. Wadami sieci neuronowych jest trudność w ustaleniu parametrów architektury, wpadanie w minima lokalne, długi czas uczenia się sieci i brak jasnej interpretacji [1,2].

Do budowy modelu prognostycznego zastosowano liniową sieć neuronową (sieć L), sieć jednokierunkową wielowarstwową ze wsteczną propagacją błędów (sieć MLP) oraz sieć o radialnych funkcjach bazowych (sieć RBF).

Sieć liniowa – jest reprezentowana przez sieć, która nie posiada warstw ukrytych, zaś znajdujące się w warstwie wyjściowej neurony są w pełni liniowe (tzn. są to neurony, w których łączne pobudzenie wyznaczone jest jako liniowa kombinacja wartości wejściowych i które posiadają liniową funkcję aktywacji).

W **sieci neuronowej jednokierunkowej wielowarstwowej ze wsteczną propagacją błędów** dobór liczby neuronów w warstwie wejściowej jest uwarunkowany wymiarem wektora danych x . Model neuronu składa się z elementu sumacyjnego, do którego dochodzą sygnały wejściowe x_1, x_2, \dots, x_N , tworzące wektor wejściowy $x = [x_1, x_2, \dots, x_N]^T$ pomnożone przez przyporządkowane im wagi $w_{i1}, w_{i2}, \dots, w_{iN}$ tworzące wektor wag i -tego neuronu $w_i = [w_{i1}, w_{i2}, \dots, w_{iN}]^T$ oraz wartość w_{i0} zwaną progiem. Sygnał wyjściowy sumatora oznaczono u_i (1), przy czym sygnał

$$u_i = \sum w_{ij} x_j + w_{i0} \quad (1)$$

jest podawany na blok realizujący nieliniową funkcję aktywacji $f(u_i)$. Najczęściej funkcja ta przyjmuje postać sigmoidalną lub tangensa hiperbolicznego. Postać funkcji stanowi ciągłe przybliżenie funkcji skokowej, przy której sygnał wyjściowy neuronu y_i przyjmuje dowolne wartości z przedziału $[0,1]$ dla funkcji unipolarnej.

Funkcja sigmoidalna unipolarna (2) to:

$$f_u(u) = \frac{1}{1 + \exp(-\beta u)}. \quad (2)$$

Parametr β pozwala regulować pochyłość sigmoidy, co bywa bardzo użyteczne.

Sieć o radialnych funkcjach bazowych (ang. *Radial Basis Function* – RBF) należy do grupy sieci warstwowych jednokierunkowych. Podstawowa różnica w stosunku do prezentowanego perceptronu wielowarstwowego leży w warstwie ukrytej, która jest w tym przypadku jedna, a neurony tej warstwy mają zaimplementowane jako *funkcje aktywacji* funkcje radialne. Najczęściej jest to *funkcja Gaussa*. Sieci radialne składają się z neuronów, których funkcje aktywacji realizują odwzorowanie (3):

$$x \mapsto \varphi(\|x - c\|), x \in R^n, \quad (3)$$

gdzie $(\|\cdot\|)$ oznacza normę euklidesową.

Funkcje $\varphi(\|x - c\|)$ nazywa się radialnymi funkcjami bazowymi. Ich wartości zmieniają się radialnie wokół centrum c .

Sieci RBF posiadają kilka zalet w porównaniu z sieciami typu MLP. Po pierwsze, jak już wcześniej stwierdzono, mogą one modelować dowolną funkcję nieliniową za pomocą pojedynczej warstwy ukrytej, przez co eliminuje się konieczność podejmowania na etapie projektowania decyzji dotyczącej liczby warstw. Po drugie prosta transformacja liniowa dokonywana w warstwie wyjściowej może być w całości zoptymalizowana przy użyciu tradycyjnych technik modelowania liniowego, które są szybkie i przy ich stosowaniu nie pojawiają się takie problemy jak minima lokalne, które występują w uczeniu sieci MLP. Z tego powodu sieci RBF mogą być uczone w bardzo krótkim czasie (różnica w szybkości uczenia dotyczy rzędów wielkości).

3. Modele przewidywania użycia narzędzi w różnych interwałach czasowych

Współcześnie w przedsiębiorstwach produkcyjnych użycie narzędzi kontroluje się bardzo ogólnie. Co określony czas sprawdzane są stany narzędzi w narzędziowniach przedsiębiorstwa.

Natomiast opracowane modele pozwolą przewidywać zapotrzebowanie na określone narzędzia w różnych interwałach czasowych i szybciej reagować na braki narzędzi w przedsiębiorstwie, co zapobiegnie przestojom produkcji i przez to również wzroste kosztów produkcji.

Modele przewidywania użycia narzędzi w różnych interwałach czasowych zostały opracowane jako modele predykcji.

Przygotowanie danych dla modeli predykcyjnych

Eksperymenty zostały przeprowadzone dla wybranych narzędzi obróbczych: frezów, noży tokarskich i ściernic. Do opracowania modeli użyto sieci neuronowych (liniowej L, MLP, RBF). Modele predykcji zostały opracowane dla przewidywanego użycia narzędzi w różnych interwałach czasowych, takich jak: godziny, dni, tygodnie i miesiące. Modele zostały nauczone poprawnego przewidywania na podstawie danych rzeczywistych w postaci przykładów użycia narzędzi do obróbki. Każdy model był testowany pod kątem prawidłowej predykcji. Wykonana została analiza i porównanie modeli.

Modele predykcji pozwoliły na przewidywanie, jakie narzędzia, do jakich operacji technologicznych były wykorzystywane najczęściej oraz w jakim czasie. Tego typu modele pozwalają na bieżące sprawdzanie dostępności narzędzi w narzędziowni oraz na generowanie zamówień zakupu narzędzi brakujących.

W celu utworzenia modeli predykcji w postaci sieci neuronowych przygotowano plik uczący i testujący. Plik uczący zawierał około 100 przykładów, a plik testowy około 12.

Dane są pogrupowane według dnia, miesiąca, czasu użycia narzędzia (tabela 1). Znając model obiektu, należy określić jego zachowanie dla różnych wymuszeń pojawiających się na wejściu tego obiektu. Interesujące jest określenie przyszłego stanu obiektu dla czasu $t + n$, gdzie n stanowi horyzont prognozy, t zawiera historię zmian wejścia do chwili bieżącej. Horyzont prognozy $n = 1$ oznacza 1 dzień. W celu zbudowania ciągów czasowych, które dalej są wykorzystywane w modelu predykcyjnym zostały dodane wartości czasu użycia narzędzia sprzed chwili t ($t - 1, t - 2, \dots, t - 7$) i po t ($t + 1$). Badania przeprowadzono również dla wartości sprzed t ($t - 30$) i po t ($t + 1$) oznaczających prognozę na miesiąc, które potwierdziły poprawność modeli.

Tabela 1. Fragment pliku uczącego [3]

		Wejście								Wyjście
Dzień	Miesiąc	Czas pracy narzędzia w chwili								
		t - 7	t - 6	t - 5	t - 4	t - 3	t - 2	t - 1	t	t + 1
1	2	30	26	45	12	34	23	45	67	11
2	2	26	45	12	34	23	45	67	11	66
3	2	45	12	34	23	45	67	11	66	34
4	2	12	34	23	45	67	11	66	34	76
5	2	34	23	45	67	11	66	34	76	23

Źródło: Badania własne.

Ocena modeli predykcji

Jakościowej oceny neuronowych modeli predykcyjnych można dokonać, porównując wykresy: rzeczywisty i prognozowany. Jest to często spotykana w literaturze metoda prezentacji wyników badań. Jednak dopiero ilościowe metody oceny modeli neuronowych umożliwiają formułowanie bardziej obiektywnych wniosków. Zasadniczo ocenę modeli neuronowych przeprowadza się dwuetapowo. W pierwszej fazie, po skonstruowaniu nauczonych sieci, do ich oceny zalecane są tzw. statystyki regresyjne. Do statystyk regresyjnych, służących do oceny modeli, po nauczaniu sieci, należą m.in. następujące metody oceny [4]:

- średnia (arytmetyczna) wyznaczona na podstawie rzeczywistych wartości prognozowanej wielkości,
- odchylenie standardowe, obliczone dla rzeczywistych danych,
- średni błąd – przeciętna wartość różnic między wartościami prognozowanymi i rzeczywistymi,
- standardowe odchylenie błędów dla zmiennej prognozowanej,
- średni błąd bezwzględny (średnia z wartości bezwzględnych różnic, wartości przewidywanych przez model oraz wartości rzeczywistych).
- iloraz odchyłeń standardowych – iloraz standardowego odchylenia błędów i odchylenia standardowego rzeczywistych danych (obu parametrów określonych wyżej),
- korelacja (właściwie standardowy współczynnik korelacji Pearsona r) – wyznaczony dla wartości rzeczywistych i wartości przewidywanych.

Największe znaczenie dla oceny jakości modeli neuronowych mają: iloraz odchyłeń standardowych oraz korelacja wartości rzeczywistych i prognozowanych [5]. Ten pierwszy parametr dla stworzonych modeli powinien przyjmować wartości rzędu 0,1...0,2. Iloraz odchyłeń o wartości bliskiej zero świadczy o dobrej jakości opracowanego modelu. Jeżeli jest on większy od jedności (lub jej bliski), to zaprojektowany model można odrzucić. Trudno jednoznacznie określić poprawność modelu, jeśli iloraz odchyłeń mieści się w przedziale: 0,3...0,7.

Ostatecznie jakość modelu zależy od uzyskanych błędów *ex post*, a ściślej mówiąc, czy są one do zaakceptowania w danym przypadku.

Korelacja wartości rzeczywistych i prognozowanych przyjmuje wartości z przedziału od 0 do 1. Najlepiej jest, gdy jest ona zbliżona do jedności (im bliżej, tym lepiej).

W drugiej fazie, po nauczaniu sieci, przeprowadza się proces prognozowania. Uzyskuje się oprócz wartości rzeczywistych, także predykcje zmiennej wyjściowej modelu. Pozwala to na wy-

znaczenie błędów ex post. Do wyliczenia błędów typu ex post należy dysponować parami wartości: rzeczywistych i wyznaczonych za pomocą danego modelu. Do najczęściej stosowanych miar jakości modeli neuronowych można zaliczyć: sumę kwadratów reszt, błąd średniokwadratowy oraz pierwiastek błędu średniokwadratowego [6]. Określone w ten sposób miary jakości błędów predykcyjnych modeli neuronowych zostały wykorzystane w fazie eksperymentów. Wartości pierwiastka błędów średniokwadratowych uczenia i testowania zostały przedstawione przy opracowanych modelach.

Został wybrany model o najmniejszym błędzie RMS oraz wskaźniku liczby wzorców poza tolerancją. Wyniki uzyskane z modeli predykcyjnych zostały porównane również z rzeczywistymi wartościami użycia narzędzi w przewidywanych przedziałach czasu.

Sieci neuronowe jako modele prognozowania

Model sieci MLP do prognozowania

Jako pierwszy zastosowano model sieci neuronowej jednokierunkowej, wielowarstwowej ze wsteczną propagacją błędów.

Zbudowane zostały różne modele sieci neuronowej. Przetestowano struktury z dwoma wejściami i jednym wyjściem oraz warstwą ukrytą zawierającą 5, 10 lub 15 neuronów. Te struktury uczono z różnymi warunkami kończącymi proces uczenia, tzn. koniec uczenia nastąpił po osiągnięciu liczby epok 1000, 10 000 lub 100 000. Porównano dla wszystkich kombinacji błąd RMS.

Najbardziej dokładnym modelem okazała się sieć neuronowa o strukturze 8-15-1, przy zakończeniu uczenia przy wartości 10 000 epok. Potwierdziły to otrzymane wyniki dla nowych wartości wejściowych. Po opracowaniu modelu sieci neuronowej dokonano jego oceny jako modelu prognostycznego.

Ewaluacja – ocena modelu neuronowego

Model prognostyczny wykonany został etapami. Etapy procedury prognostycznej są następujące:

- obserwacja procesu w postaci ciągu czasowego,
- utworzenie modelu obserwowanego ciągu,
- przekształcenie modelu ciągu czasowego do postaci umożliwiającej prognozę (budowa algorytmu predykcji).

Do prognozy czasu pracy narzędzia wybrano tryb prognoz jednoseryjnych dla modeli ciągów czasowych. Pierwszy etap to przedstawienie ciągu czasowego zawierającego dane dotyczące czasu pracy narzędzia. Następnie został użyty model sieci neuronowej do prognozowania użycia narzędzia. Ocena dokładności prognozy możliwa jest w trybie testowania. W trybie tym pewną część wejściowego ciągu czasowego wyłączamy ze zbioru, przy pomocy którego identyfikowane są parametry modelu. Tryb testowania wykazał dobre dopasowanie modelu dla danych historycznych, na podstawie których wykonywana jest prognoza.

Kolejny etap badań obejmuje modele predykcyjne do przewidywania użycia narzędzi opracowane przy pomocy innych typów sieci neuronowych oraz analiza przydatności różnych typów sieci do budowania modeli predykcyjnych.

Model liniowej sieci neuronowej do prognozowania

Podobne badania wykonano dla sieci liniowej. Najbardziej dokładnym modelem okazała się sieć neuronowa o strukturze 8-1, przy zakończeniu uczenia przy wartości 10 000 epok. Potwierdzi-

ły to otrzymane wyniki dla nowych wartości wejściowych. Jednak jest to model dużo gorszy od modelu opracowanego przy użyciu sieci MLP.

Model sieci RBF do prognozowania

W tabeli 2 pokazano parametry modeli predykcji z użyciem sieci o radialnych funkcjach bazowych (RBF).

Tabela 2. Parametry sieci RBF

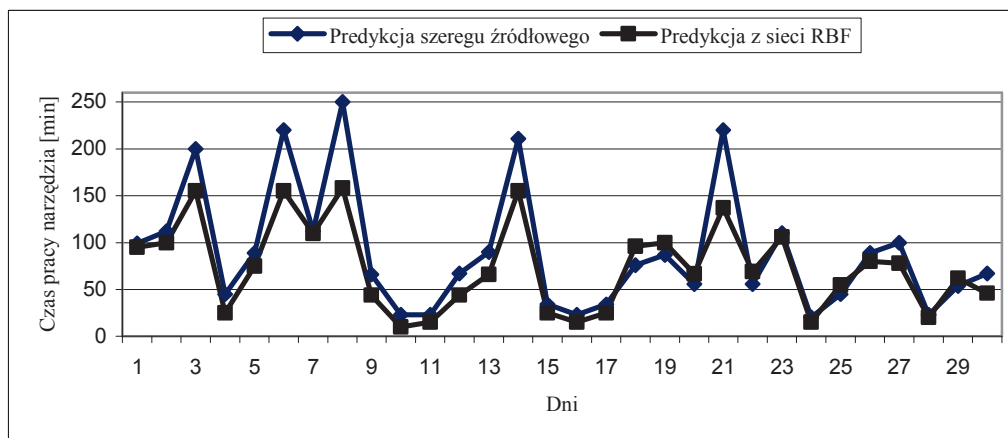
Typ sieci	Jakość uczenia	Jakość testowania	Błąd uczenia	Błąd testowania	Liczba wejść	Liczba neuronów w warstwie ukrytej
RBF 8-7-1	0,9295	0,9089	0,2159	0,4218	8	7
RBF 8-9-1	0,9308	0,9115	0,2127	0,4140	8	9
RBF 8-15-1	0,9602	0,9421	0,1832	0,3585	8	15
RBF 8-20-1	0,9895	0,9687	0,1563	0,3213	8	20

Źródło: Badania własne.

Na rysunku 1 pokazano predykcję szeregu źródłowego i predykcję z sieci RBF.

Porównanie prostych modeli prognostycznych

W tabeli 3 przedstawiono porównanie modeli prognostycznych przy użyciu sieci neuronowych liniowych, sieci MLP oraz sieci RBF. Przetestowano te modele dla wybranych narzędzi obróbczych. Nie zauważono potrzeby różnicowania modeli względem różnych narzędzi. Najbardziej dokładnym modelem okazał się model oparty na sieci RBF.



Rysunek 1. Prognoza przy użyciu sieci neuronowej RBF

Źródło: Badania własne.

Tabela 3. Porównanie modeli prognostycznych przy użyciu sieci neuronowych liniowych, MLP oraz RBF

Typ sieci	Jakość uczenia	Jakość testowania	Błąd uczenia	Błąd testowania	Liczba wejść	Liczba neuronów w warstwie ukrytej
MLP 8-15-1	0,9312	0,9156	0,1599	0,3786	8	15
Liniowa 8-1	0,6489	0,6325	0,5581	0,5943	8	0
RBF 8-20-1	0,9895	0,9687	0,1563	0,3213	8	20

Źródło: Badania własne.

4. Podsumowanie

Zastosowanie sieci neuronowych we wspomaganie projektowania procesów technologicznych, szczególnie w przewidywaniu użycia narzędzi w określonych interwałach czasowych wprowadziło nową jakość do systemów CAPP oraz może stać się podstawą algorytmizacji nowych systemów zwanych „inteligentnymi”.

Wykorzystanie metod sztucznej inteligencji umożliwia utworzenie systemu wspomaganie, który w sposób automatyczny pozyskuje wiedzę i ma właściwość adaptacji. Jest to szczególnie ważne przy opracowywaniu systemu dla złożonych systemów rzeczywistych, w których następują ciągle zmiany i jedne podprocesy zależą od innych oraz wiele czynników zależy jeden od drugiego i każda zmiana powoduje zmiany kolejne.

Korzyści ze stosowania metod sztucznej inteligencji wielokrotnie przekraczają pracochłonność ich tworzenia. W rzeczywistości najdłuższym etapem procesu ich tworzenia jest zgromadzenie i przygotowanie danych źródłowych, chociaż poprzez automatyczne gromadzenie danych również ten etap ulega skróceniu.

Modele przewidywania użycia narzędzi są wykorzystane już na etapie projektowania procesu technologicznego, aby w odpowiednim czasie zostały wytworzone wyroby i nie doszło do przestoju z powodu braku narzędzi, szczególnie narzędzi specjalnych.

Bibliografia

- [1] Grzeszczyk T. A., Sztuczna inteligencja we wspomaganie procesu prognozowania w przedsiębiorstwie, StatSoft Polska, s. 91–105, Kraków, 2005.
- [2] Krawiec K., Stefanowski J., Uczenie maszynowe i sieci neuronowe, wydawnictwo Politechniki Poznańskiej, Poznań, 2004.
- [3] Lula P., Ocena modeli neuronowych wykorzystywanych w zagadnieniach modelowania i prognozowania, w: Jajuga K., Walesiak M. (red.), Klasyfikacja i analiza danych – teoria i zastosowania, Prace Naukowe Akademii Ekonomicznej, Nr 874, Taksonomia 7, s. 129–138, Wrocław, 2000.
- [4] Rojek I., Wspomaganie procesów podejmowania decyzji i sterowania w systemach o różnej skali złożoności z udziałem metod sztucznej inteligencji, wydawnictwo UKW, Bydgoszcz, 2010.
- [5] Rutkowski L., Metody i techniki sztucznej inteligencji, Inteligencja obliczeniowa, Wydawnictwa Naukowe PWN, Warszawa, 2005.

- [6] Tadeusiewicz R., Lula P., *Statistica Neural Networks 4.0 PL: Wprowadzenie do sieci neuronowych*, StatSoft Polska, Kraków, 2001.

NEURAL NETWORKS AS FORECASTING MODELS OF TOOL USE IN CAPP SYSTEMS

Summary

In the paper the neural networks were presented as the forecasting models of tool use in CAPP systems. The models were worked out by use linear neural networks, multi-layer networks with error back propagation and Radial Basis Function networks.

The comparison of these models was executed. The effectiveness of forecasting of tool use in different time intervals is the measure of model evaluation. These models are used on the stage of design of manufacturing process in aim planned producing products and that it did not come to standstill with reason of lack of tools, particularly special tools. Created models were tested on real data from the enterprise.

Keywords: neural network, forecasting model, tool

Izabela Rojek
Instytut Mechaniki Środowiska i Informatyki Stosowanej
Uniwersytet Kazimierza Wielkiego
ul. Chodkiewicza 30, 85-064 Bydgoszcz
e-mail: izarojek@ukw.edu.pl