

Bogusław ŁAZARZ, Piotr CZECH

OPTIMALIZACJA SIECI NEURONOWEJ TYPU SVM W ZADANIU IDENTYFIKACJI STOPNIA PĘKNIĘCIA PODSTAWY ZĘBA

Streszczenie. W opracowaniu przedstawiono wyniki eksperymentu, którego celem był najlepszy z możliwych dobór parametrów sztucznej sieci neuronowej typu SVM w zadaniu klasyfikacji stopnia pęknięcia podstawy zęba.

OPTIMIZATION OF THE SVM NEURAL NETWORK IN TASK OF THE DEGREE OF TOOTH ROOT CRACKING IDENTIFICATION SYSTEM

Summary. The work presents results of an experiment that employs the artificial neuronal network in the task of identification of the degree of tooth root cracking.

1. WPROWADZENIE

W dzisiejszych czasach przekładnie zębate należą do najbardziej rozpowszechnionych podzespołów wszelkich maszyn i zespołów maszyn. Bardzo ważne w czasie eksploatacji jest umiejętność diagnozowania i odpowiednio wczesne wykrywanie uszkodzeń przekładni [9,12]. Do najczęściej stosowanych metod należą badania wibroakustyczne. Ich największą zaletą jest to, że mogą być przeprowadzane bez demontażu maszyny. Cały czas trwają poszukiwania metody, która będzie w stanie pewnie i szybko wykryć uszkodzenie i to nie tylko w jego zaawansowanym stadium, ale także w stadiach początkowych.

Rozwijające się obecnie bardzo intensywnie sztuczne sieci neuronowe wydają się szansą na jego rozwiązanie [1]. Możliwości sztucznej inteligencji, do których należą zdolność uczenia się i przewidywanie na podstawie zdobytej wiedzy, pozwalają sądzić, że w niedalekiej przyszłości systemy analizy danych, w których elementem analizującym nie będzie człowiek,

a właśnie jakiś rodzaj sztucznej inteligencji

Zasada działania sztucznych sieci neuronowych oparta jest na zdolności uczenia się [7,11]. W procesie uczenia sieci neuronowe nabywają umiejętności generalizacji nabytej wiedzy [11]. Proces uczenia sieci neuronowych może odbywać się z nauczycielem lub bez (sieci samoorganizujące) [7,11].

Sztuczna sieć neuronowa składa się z połączonych ze sobą neuronów, budową przypominając naturalny mózg ludzki [7,11]. Do każdego neuronu docierają dane, podawane na jego wejście z odpowiednią „siłą” (wagą). Proces uczenia sztucznej sieci neuronowej

polega na ustaleniu wag poszczególnych neuronów [7,11]. Sztuczne sieci neuronowe mają tendencję do „samooczyszczania”, czyli do samoistnej likwidacji połączeń neuronowych, jeśli nie są przydatne w rozwiązywaniu problemu.

W literaturze można spotkać się z wieloma typami sztucznych sieci neuronowych, używanych w zadaniach klasyfikacji, klasteryzacji, aproksymacji, prognozowania [1,7,11].

2. OPIS EKSPERYMENTU

Celem eksperymentu była próba użycia sztucznej sieci neuronowej w problemie oceny stopnia pęknięcia u podstawy zęba. W literaturze można spotkać udane eksperymenty mające na celu wykorzystanie metod sztucznej inteligencji w diagnostyce stanu maszyn [1].

Za obiekt badań posłużyła przekładnia zębata pracująca w układzie mocy krążącej.

W eksperymencie zidentyfikowany model przekładni zębatej pracującej w układzie mocy krążącej [2] posłużył do symulacji przyspieszeń drgań zębniaka przy różnym stopniu pęknięcia u podstawy zęba, w zakresie 0 do 100%. Symulowanie pęknięcia zęba u podstawy polegało na obniżeniu jego sztywności o ustaloną wartość w stosunku do sztywności zęba nieuszkodzonego. Symulacje zostały przeprowadzone dla różnych klas dokładności wykonania kół zębatych oraz dla różnych wartości błędów cyklicznych i losowych. Dla kół zębatych w każdej przyjętej klasie dokładności przeprowadzono symulacje dla uszkodzenia zęba w postaci podcięcia jego podstawy w zakresie od 0 do 100%. W sumie otrzymano 1111 symulacji.

Następnie uzyskane z symulacji sygnały poddano dalszej obróbce uzyskując sygnały czasowe:

- z widma przyspieszeń drgań w całym zakresie częstotliwości (nr porz. 1),
- z widma przyspieszeń drgań w zakresie od 0 do 6 kHz (nr porz. 2),
- z widma przyspieszeń drgań w zakresie od 0 do 6 kHz sygnału różnicowego (nr porz. 3),
- z widma przyspieszeń drgań w zakresie od 0 do 6 kHz z sygnału resztkowego (nr porz. 4),
- z widma przyspieszeń drgań w zakresie od $\frac{1}{2}f_z$ do $\frac{3}{2}f_z$ (nr porz. 5),
- z widma obwiedni przyspieszeń drgań w całym zakresie częstotliwości (nr porz. 6),
- z widma obwiedni przyspieszeń drgań w zakresie od 0 do 6 kHz (nr porz. 7),
- z widma obwiedni przyspieszeń drgań w zakresie od 0 do 6 kHz sygnału różnicowego (nr porz. 8),
- z widma obwiedni przyspieszeń drgań w zakresie od 0 do 6 kHz z sygnału resztkowego (nr porz. 9),
- z widma obwiedni przyspieszeń drgań w zakresie od $\frac{1}{2}f_z$ do $\frac{3}{2}f_z$ (nr porz. 10),
- z widma poliharmonicznego przyspieszeń drgań w zakresie od 0 do 6 kHz (nr porz. 11),
- z widma poliharmonicznego przyspieszeń drgań w zakresie od 0 do 6 kHz sygnału różnicowego (nr porz. 12),
- z widma poliharmonicznego przyspieszeń drgań w zakresie od 0 do 6 kHz z sygnału resztkowego (nr porz. 13),
- z widma iloczynowego przyspieszeń drgań w zakresie od 0 do 6 kHz (nr porz. 14),
- z widma iloczynowego przyspieszeń drgań w zakresie od 0 do 6 kHz sygnału różnicowego (nr porz. 15),

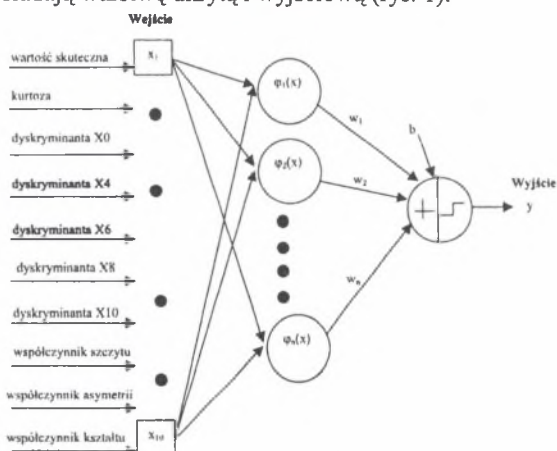
- z widma iloczynowego przyspieszeń drgań w zakresie od 0 do 6 kHz z sygnału resztkowego (nr porz. 16),
- korelacji sygnału czasowego uzyskanego z widma przyspieszeń drgań w zakresie od 0 do 6 kHz (nr porz. 17),
- korelacji sygnału czasowego uzyskanego z widma przyspieszeń drgań w zakresie od 0 do 6 kHz sygnału różnicowego (nr porz. 18),
- korelacji sygnału czasowego uzyskanego z widma przyspieszeń drgań w zakresie od 0 do 6 kHz z sygnału resztkowego (nr porz. 19),
- kowariancji sygnału czasowego uzyskanego z widma przyspieszeń drgań w zakresie od 0 do 6 kHz (nr porz. 20),
- kowariancji sygnału czasowego uzyskanego z widma przyspieszeń drgań w zakresie od 0 do 6 kHz sygnału różnicowego (nr porz. 21),
- kowariancji sygnału czasowego uzyskanego z widma przyspieszeń drgań w zakresie od 0 do 6 kHz z sygnału resztkowego (nr porz. 22).

Dla uzyskanych w ten sposób sygnałów czasowych obliczono miary statystyczne [3,5,9,10].

Spośród szeregu miar za dane wejściowe do sztucznej sieci neuronowej wybrano następujące miary statystyczne:

- wartość skuteczną,
- kurtozę,
- dyskryminantę X0,
- dyskryminantę X4,
- dyskryminantę X6,
- dyskryminantę X8,
- dyskryminantę X10,
- współczynnik szczytu,
- współczynnik asymetrii,
- współczynnik kształtu.

Za typ sztucznej sieci neuronowej posłużyła sieć SVM (Support Vector Machine) [4,6,8]. Sieć ta należy do sieci jednokierunkowych. Sieci te mają najczęściej budowę dwuwarstwową i posiadają warstwę ukrytą i wyjściową (rys. 1).



Rys. 1. Struktura sztucznej sieci neuronowej typu SVM

Fig. 1. Structure of SVM Neural Network

Za funkcję aktywacji $\varphi(x)$ w tym typie sieci można przyjąć:

- funkcję liniową,
- funkcję wielomianową,
- funkcję radialną,
- funkcję sigmoidalną.

W eksperymencie badawczym za funkcję aktywacji przyjęto funkcję radialną. Sieci neuronowe typu SVM wymagają uczenia z nauczycielem [6,8].

Problem klasyfikacji wzorców nie separowalnych liniowo sprowadza się do określenia takiej optymalnej hiperpłaszczyzny, która zminimalizuje prawdopodobieństwo wystąpienia błędu klasyfikacji na zbiorze uczącym z możliwie najszerszym marginesem separacji.

W eksperymencie podzielono zakres stopnia pęknięcia podstawy zęba na pięć klas: $0 \div 20\%$, $21 \div 40\%$, $41 \div 60\%$, $61 \div 80\%$, $81 \div 100\%$ do których nauczona sztuczna sieć neuronowa miała klasyfikować przypadki z uszkodzeniami kół zębatych.

W celu zmniejszenia błędu klasyfikacji zmieniano następujące współczynniki sieci SVM:

- szerokości marginesu błędu uczenia i testowania sieci (ϵ),
- kosztu (C),
- wartości gamma.

Szerokość marginesu błędu określa dopuszczalną odchyłkę, dla której wyniki o mniejszej odchyłce nie są traktowane jako błąd.

Parametr kosztu jest parametrem sterującym złożonością sieci SVM. Określa on wagę, z jaką traktujemy błędy testowania sieci w stosunku do ustalonego marginesu separacji.

Parametr gamma określony jest dla radialnej funkcji jądra z zależności: $\gamma = \frac{1}{\sigma^2}$.

Podczas badań optymalizowano parametry sieci SVM poszukując najmniejszej wartości błędu klasyfikacji dla stałej wartości współczynników ϵ i γ zmieniając wartość współczynnika C. Następnie dla sieci charakteryzującej się najmniejszym błędem klasyfikacji dla ustalonego współczynnika C poszukiwano wartość współczynnika γ , dla której sieć popełniła najmniej błędów.

Algorytm ten został zastosowany do sygnałów dla każdego z 22 przebiegów czasowych reprezentowanych przez wybrane 10 miar statystycznych.

3. WYNIKI I WNIOSKI EKSPERYMENTU BADAWCZEGO

Wyniki eksperymentu zostały przedstawione w tabeli 1. Składa się ona z dziewięciu kolumn, w których kolejno przedstawiono numer porządkowy obróbki sygnału przyspieszeń drgań, wartości współczynników ϵ , γ , C, złożoność sieci (SV), ilość błędów w czasie procesu uczenia sieci i jej procentową wartość, ilość błędów w czasie procesu testowania sieci i jej procentową wartość.

Na podstawie wyników można zauważyć zależność pomiędzy złożonością sieci neuronowej, czyli ilością wektorów podtrzymujących (SV), a ilością popełnianych przez sieć błędów w czasie testowania. Jest to zależność odwrotnie proporcjonalna.

Najmniejszy procent błędów testowania uzyskano dla kowariancji sygnału różnicowego (czasowego), który powstał z widma przyspieszeń drgań w zakresie od 0 do 6 kHz (nr porz. 21). Błąd ten wynosi 7,2072% dla sieci o 227 wektorach podtrzymujących. Równocześnie sieć ta uzyskiwała wartość błędu uczenia na poziomie 4,1367%.

Tabela 1

Wyniki eksperymentu

Nr porz.	ϵ	γ	C	SV	Błąd uczenia		Błąd testowania	
					ilość	%	ilość	%
1	0,1	20	52000	313	64	11,5108	84	15,1351
2	0,1	10	66000	313	89	16,0072	105	18,9189
3	0,1	20	77000	308	24	4,3165	76	13,6937
4	0,1	3	90000	306	85	15,2878	109	19,6396
5	0,1	10	20000	391	13	2,3381	216	38,9189
6	0,1	70	30000	356	57	10,2518	98	17,6577
7	0,1	100	2000	436	89	16,0072	112	20,1802
8	0,1	20	30000	262	26	4,6763	52	9,3694
9	0,1	30	6000	302	34	6,1151	61	10,991
10	0,1	10	3000	364	63	11,3309	177	31,8919
11	0,1	40	30000	261	60	10,7914	71	12,7928
12	0,1	80	30000	225	40	7,3741	61	10,991
13	0,1	100	18000	233	35	6,295	70	12,6125
14	0,1	50	30	552	6	1,8791	409	73,6937
15	0,1	50	3	552	7	1,259	414	74,5946
16	0,1	0,01	3	546	197	35,4317	385	69,3694
17	0,1	30000	600	292	9	1,6187	47	8,4685
18	0,1	5000	300	282	18	3,2374	46	8,2883
19	0,1	100	15000	247	40	7,1942	74	13,3337
20	0,1	3000	300	360	30	5,3957	77	14,2342
21	0,1	1000	30000	227	23	4,1367	40	7,2072
22	0,1	100	30000	301	43	7,7338	101	18,1982

Z badań wynika również, iż wielkość błędu uczenia maleje, gdy błąd testowania zaczyna już rosnąć. Wynika to z faktu, że sieć zaczyna się uczyć na pamięć (przeuczenia sieci). Stąd też wynika fakt uzyskania znacznie lepszych wyników dla procesu uczenia niż późniejszej weryfikacji w procesie testowania sieci.

Dalsze badania mogą objąć poszukiwanie odpowiednich miar sygnałów uzyskanych z cepstrum, bispectrum, transformaty falkowej lub Wignera-Ville'a, lub zastosowania innego rodzaju sztucznej sieci neuronowej.

Literatura

1. Korbicz J., Kościelny J., Kowalczyk Z., Cholewa W.: Diagnostyka procesów. Modele. Metody sztucznej inteligencji. Zastosowania, Wydawnictwo Naukowo-Techniczne, Warszawa 2002.
2. Łazarz B.: Zidentyfikowany model dynamiczny przekładni zębatej jako podstawa projektowania. Biblioteka Problemów Eksploatacji, Katowice-Radom 2001.
3. Madej H., Czech P., Konieczny Ł.: Diagnostyka przekładni zębatych klatek walcowniczych, Diagnostyka nr 28, 2003.
4. Mangasarian O. L.: Lagrangian Support Vector Machines, Journal of Machine Learning Research, 2001.

5. Martin H. R.: Statistical Moment Analysis as a Means of Surface Damage Detection. Proceedings of the 7th International Modal Analysis Conference, Society of Experimental Mechanics, Schenectady, New York 1989.
6. Osowski S.: Sieci neuronowe SVM w zastosowaniu do problemów regresji, Przegląd Elektrotechniczny nr 86, 2002.
7. Osowski S.: Sieci neuronowe do przetwarzania informacji, Oficyna Wydawnicza Politechniki Warszawskiej, Warszawa 2000.
8. Osowski S.: Zastosowanie sieci SVM do klasyfikacji wzorców, Przegląd Elektrotechniczny nr 78, 2002.
9. Radkowski S.: Wibroakustyczna diagnostyka uszkodzeń niskoenergetycznych, Biblioteka Problemów Eksploatacyjnych, Warszawa-Radom 2002.
10. Stewart R. M.: Some Useful Data Analysis Techniques for Gearbox Diagnostics. Report MHM/R/10/77, Machine Health Monitoring Group, Institute of Sound and Vibration Research, University of Southampton 1977.
11. Tadeusiewicz R.: Sieci neuronowe, Akademicka Oficyna Wydawnicza, Warszawa 1993.
12. Wilk A., Łazarz B., Madej H.: Vibration Processing Techniques for Fault Detection in Gearboxes, Proceedings of DETC'2003 ASME 2003 Design Engineering Technical Conferences and Computers and Information in Engineering Conference Chicago, Illinois, USA, 2003.

Abstract

Results of the experiment show inversely proportional dependence of the network complexity, i. e. on amount of the SV parameters, upon the defect range. Minimal percent of the test defects equals 7,2072 % for network of 227 SV. The experiment indicates to necessity of further studies on the problem in order to find such measures that can give a detailed diagnostic description of the tooth damage.

Recenzent: Prof. dr hab. inż. Tadeusz Wieczorek