



SZKOŁA GŁÓWNA GOSPODARSTWA WIEJSKIEGO W WARSZAWIE

Wybrane problemy predykcji i klasyfikacji sieci SVM w elektrotechnice i bioinżynierii

Autoreferat

Robert SAŁAT

**Katedra Podstaw Inżynierii
Wydział Inżynierii Produkcji
Szkoła Główna Gospodarstwa Wiejskiego**

Warszawa 2015

- I. Imię i nazwisko. Posiadane stopnie naukowe z podaniem nazwy, miejsca i roku ich uzyskania oraz tytułu rozprawy doktorskiej

Robert Sałat

Doktor nauk technicznych w zakresie elektrotechniki, nadany przez Radę Wydziału Elektrycznego Politechniki Warszawskiej w dniu 11.12.2002 na podstawie rozprawy doktorskiej zatytułowanej „Zastosowanie sieci neuronowych do diagnostyki układów elektrycznych i elektroenergetycznych”.

Promotor: Prof. dr hab. Stanisław Osowski

Recenzenci: Prof. dr hab. Jan Machowski, Prof. dr hab. Tadeusz Łobos

- II. Informacja o dotychczasowym zatrudnieniu w jednostkach naukowych

Od października 2002 roku do chwili obecnej jestem zatrudniony w Katedrze Podstaw Inżynierii Wydziału Inżynierii Produkcji Szkoły Głównej Gospodarstwa Wiejskiego w Warszawie; od października 2003 roku na etacie adiunkta.

- III. Wskazanie osiągnięcia naukowego wynikającego z art. 16 ust. 2 ustawy z dnia 14.03.2003 o stopniach naukowych i tytule naukowym

1. Tytuł osiągnięcia naukowego

Osiągnięciem naukowym będącym podstawą habilitacji jest cykl publikacji powiązanych tematycznie (dziewięć publikacji) pod wspólną nazwą „Wybrane problemy predykcji i klasyfikacji sieci SVM w elektrotechnice i bioinżynierii”.

2. Wykaz wybranego cyklu publikacji (autorzy, tytuły publikacji, nazwa czasopisma, rok wydania, nazwa wydawnictwa, wskaźnik Impact Factor i punktacja MNiSW)

[1] Sałat R., Osowski S., Accurate fault location in the power transmission line using Support Vector Machine approach. **IEEE Transactions on Power Systems**, 2/2004, 979–986. (IEEE; IF₂₀₀₄ 0.821; MNiSW₂₀₀₄ 20 pkt.) (IF₂₀₁₄ 3.530; MNiSW₂₀₁₄ 45 pkt.)

[2] Sałat R., Osowski S., Support Vector Machine for soft fault location in electrical circuits. **Journal of Intelligent and Fuzzy Systems**, 22(1)/2011, 21–31. (IOS Press; IF₂₀₁₁ 0.556; MNiSW₂₀₁₁ 15 pkt.) (IF₂₀₁₄ 0.936; MNiSW₂₀₁₄ 15 pkt.)

- [3] Wrotek H., **Sałat R.**, Wyznaczenie parametrów ruchowych silnika reluktancyjnego przełączalnego z toczącym się wirnikiem przy wykorzystaniu metody elementu skończonego i sztucznych sieci neuronowych. **Przegląd Elektrotechniczny**, 12/2010, 276–279. (**SIGMA-NOT; IF₂₀₁₀ 0.242; MNiSW₂₀₁₀ 15 pkt.**) (**IF₂₀₁₄ 0; MNiSW₂₀₁₄ 10 pkt.**)
- [4] **Sałat R.**, Awtoniuk M., Odwzorowanie dynamiki pracy regulatora PID zaimplementowanego w sterowniku PLC za pomocą Least Squares-Support Vector Machines. **Przegląd Elektrotechniczny**, 5a/2012, 227-230. (**SIGMA-NOT; IF₂₀₁₂ 0.243; MNiSW₂₀₁₂ 15 pkt.**) (**IF₂₀₁₄ 0; MNiSW₂₀₁₄ 10 pkt.**)
- [5] **Salat R.**, Awtoniuk M., Black-Box modeling of PIDs implemented in PLCs without structural information: A Support Vector Regression approach. **Neural Computing and Applications**, 2014, DOI: 10.1007/s00521-014-1754-2. (**Springer; IF₂₀₁₄ 1.763; MNiSW₂₀₁₄ 25 pkt.**)
- [6] **Salat R.**, Salat K., New approach to predicting proconvulsant activity with the use of Support Vector Regression. **Computers in Biology and Medicine**, 42/2012, 575–581. (**Elsevier; IF₂₀₁₂ 1.162; MNiSW₂₀₁₂ 20 pkt.**) (**IF₂₀₁₄ 1.475; MNiSW₂₀₁₄ 25 pkt.**)
- [7] **Sałat R.**, Sałat K., The application of support vector regression for prediction of the antiallodynamic effect of drug combinations in the mouse model of streptozocin-induced diabetic neuropathy. **Computer Methods and Programs in Biomedicine**, 111(2)/2013, 330–337. (**Elsevier; IF₂₀₁₃ 1.093; MNiSW₂₀₁₃ 25 pkt.**) (**IF₂₀₁₄ 1.093; MNiSW₂₀₁₄ 25 pkt.**)
- [8] **Salat R.**, Salat K., Modeling analgesic drug interactions using support vector regression: A new approach to isobolographic analysis. **Journal of Pharmacological and Toxicological Methods**, 71/2015, 95-102. (**Elsevier; IF₂₀₁₄ 2.15; MNiSW₂₀₁₄ 25 pkt.**)
- [9] Winiczenko R., **Salat R.**, Awtoniuk M., Estimation of tensile strength of ductile iron friction welded joints using hybrid intelligent methods. **Transactions of Nonferrous Metals Society of China**, 23(2)/2013, 385–391. (**Elsevier; IF₂₀₁₃ 1.001; MNiSW₂₀₁₃ 30 pkt.**) (**IF₂₀₁₄ 1.001; MNiSW₂₀₁₄ 35 pkt.**)

- Sumaryczny *Impact Factor* cyklu publikacji: 9.031/11.948*
- Sumaryczna ilość punktów MNiSW dla cyklu publikacji: 190/215*
- Średni udział % w cyklu publikacji: 67%
- Indeks Hirscha dorobku naukowego wg *Web of Science* na dzień 18.02.2015: 5**
- łączna liczba cytowań wg bazy ISI Web of Science na dzień 18.02.2015: 87**
- łączny *Impact Factor* dorobku naukowego: 16.517/19.753*
- łączna liczba punktów MNiSW dla dorobku naukowego po doktoracie: 331/367*
- łączna liczba artykułów indeksowanych w bazie Journal Citation Reports (JCR): 13
- łączna liczba artykułów w czasopismach naukowych: 26
- łączna liczba artykułów (ogólnie): 47
- Podręczniki akademickie i rozdziały w monografiach naukowych: 4

*Liczby z gwiazdką odnoszą się do wyliczeń z roku 2014

** Uwzględniono cytowania, które nie są jeszcze obecne na stronie *Web of Science* ze względu na opóźnienia uaktualniania sięgające nawet do sześciu miesięcy

3. Oświadczenia współautorów cyklu publikacji

4. Omówienie celu naukowego ww. prac i osiągniętych wyników wraz z omówieniem ich ewentualnego wykorzystania

Przedstawiony do oceny cykl publikacji składa się z dziewięciu prac opublikowanych w latach 2004-2014. Dotyczy on tematu „**Wybrane problemy predykcji i klasyfikacji sieci SVM w elektrotechnice i bioinżynierii**”.

We wszystkich tych pracach, oprócz prac [3, 9], jestem pierwszym autorem; w sześciu pracach również autorem korespondencyjnym. Prace te zostały opublikowane w czasopiśmie o zasięgu światowym, indeksowanych w JCR.

4.1 WSTĘP

Sieci wektorów nośnych (SVM) są dobrze znanymi technikami uczenia z nauczycielem opartymi na teorii metod statystycznych (SLT), które są często wykorzystywane w zadaniach predykcji i klasyfikacji. W przypadku poprawnego procesu uczenia, sieci SVM cechuje bardzo duża dokładność i dobre właściwości generalizacyjne. Z wykorzystaniem sieci SVM jako predyktora lub klasyfikatora wiążą się jednak istotne problemy, których rozwiązanie jest zadaniem bardzo trudnym. Nie ma gotowej procedury na ich wykorzystanie i do każdego zadania wymagane jest podejście indywidualne. Na rozwiązanie innych problemów należy zwrócić uwagę w zadaniach predykcji a na inne w klasyfikacji. Istnieje jednak pewna grupa problemów, które należy rozwiązać zarówno w zadaniach predykcji jak i klasyfikacji.

Jednym z głównych jest dobór wektora cech¹ (ang. *feature selection*), czyli sygnałów wejściowych podawanych na sieć. Problem ten jest szczególnie istotny z uwagi na fakt, iż nie zawsze wszystkie cechy niosą istotną informację. Część z nich może być nieistotna dla rozwiązywanego zadania, część może być szumem, inna część skorelowana lub bez znaczenia dla procesu. Wybór optymalnego wektora cech sprowadza się zatem do znalezienia takiego ich zestawu, który jest istotny dla danego zadania. Chociaż dobór optymalnego zestawu cech jest zadaniem trudnym, jego rozwiązanie wnosi wiele korzyści manifestujących się, jako:

- ograniczenie zbioru danych, co pozwala uniknąć tzw. przekleństwa wielkowymiarowości (ang. *curse of dimensionality*),
- uzyskanie lepszej dokładności estymacji,
- zastosowanie bardziej skutecznych technik uczenia,
- polepszenie zdolności generalizacji sieci, może być osiągnięte, dzięki usunięciu nieistotnych cech interferujących ze sobą.

¹ Jako cechy w tym przypadku należy rozumieć dane, które mogą być otrzymane bezpośrednio z pomiarów, jak i przetworzone.

Większość metod selekcji cech polega na tzw. procedurach przeszukania, które dokonują eksploracji na całym zestawie danych. Ogólnie metody selekcji cech można podzielić na dwie grupy:

- metody niezależne - tzw. filtry (ang. *filters methods*), gdzie selekcji cech dokonuje się niezależnie od uczenia sieci. Często wykorzystywane są tu analizy głównych składowych (PCA), ich rozszerzenie o nieliniowość (KPCA), analiza składowych niezależnych (ICA) i wiele innych, których ideą jest przekształcenie oryginalnego zestawu cech w nowy nieskorelowany zestaw o znacznie mniejszym wymiarze. Do selekcji cech, szczególnie w identyfikacji, używa się również metod statystycznych np. analizy wariancji (ANOVA), metody skurczania i selekcji (LASSO), itp. Wadą tego typu metod może być utrata informacji podczas dwuetapowej procedury selekcji.
- metody połączone z uczeniem sieci SVM (ang. *wrappers and embedded methods*), gdzie dobór wektora cech jest natychmiast sprawdzany podczas uczenia i testowania. Do grupy metod tego typu można zaliczyć na przykład metodę opartą o model nieliniowej autoregresji z sygnałem zewnętrznym (NARX) oraz różne jej odmiany. Metody tego typu dają z reguły dokładniejsze rezultaty niż metody niezależne, ale ich wadą jest konieczność wykonania dużej liczby obliczeń związanych z wielokrotnym trenowaniem sieci.

Z doborem (redukcją) wektora cech należy być ostrożnym, gdyż dla problemów o dużym wymiarze wektora mogą pojawić się przypadki, gdzie wszystkie cechy są równie ważne i są tego samego typu, a redukcja ich liczby wiąże się ze znacznym pogorszeniem wyników klasyfikacji. Istnieją również przypadki, gdzie cechy, które indywidualnie uważane są za nieistotne, w połączeniu z innymi cechami, znacząco poprawiają wyniki klasyfikacji lub predykcji. Niekiedy nawet połączenie razem cech, mało istotnych indywidualnie powoduje, że stają się one bardzo użyteczne w rozwiązaniu zadania.

Problem selekcji wektora cech nie omija również zadań gdzie wektor cech jest na tyle mały, że jego redukcja jest niewskazana a nawet często wymagane jest jego rozszerzenia. Główny nacisk kładzie się wtedy na dobre zrozumienie zjawiska i zastosowaniu zaawansowanych metod doboru optymalnych cech. Szczególnie istotne jest to w bioinżynierii, gdzie dąży się aby cechy były przedstawione jako czynniki, które na przykład wywołują poszczególne choroby lub ujawniają łańcuch zdarzeń od przypadków do symptomów.

Równie istotnym problemem pojawiającym się wraz z wykorzystaniem sieci SVM, jest selekcja modelu, czyli wybór jądra (ang. *kernel selection*) oraz związany z tym dobór parametrów samej sieci (ang. *tuning parameters*), co ma istotny wpływ na wynik uczenia i testowania. Niekiedy dobór parametrów i konstrukcji samej sieci SVM jest ściśle powiązany z metodą selekcji cech, co pomaga w wyznaczeniu optymalnych parametrów sieci. W przypadku małej liczby danych, co jest dość częstym problemem występującym w bioinżynierii, bardzo popularne są różne odmiany metody n-krotnej walidacji krzyżowej (ang. *n-cross validation*) lub strategii „pozostaw jedną do testowania” (ang. *leave one out*), które w połączeniu z metodą przeszukania przestrzeni parametrów (ang. *grid search*)

prowadzą do selekcji modelu i doboru parametrów, które zapobiegają występowaniu tzw. przeuczenia sieci (ang. *overfitting*). Jeżeli danych jest odpowiednio dużo, co często występuje w zadaniach elektrotechniki, w wyznaczeniu optymalnych parametrów sieci SVM można wykorzystać inne techniki sztucznej inteligencji, jak na przykład algorytmy genetyczne.

Z selekcją modelu ściśle powiązana jest strategia uczenia, która jest różna dla zadania klasyfikacji i predykcji. Szczególnie w zadaniach klasyfikacji, gdzie liczba klas może być dowolna, bardzo ważny jest wybór odpowiedniej metody, gdyż ze swej natury sieci SVM dokonują rozdziału danych na dwie klasy. Zwiększenie rozpoznania wielu klas wymaga na przykład zastosowania metod typu „jeden przeciw pozostałym” (ang. *one versus all*) lub „jeden przeciw jednemu” (ang. *one versus one*). Niedogodnością tych metod jest potrzeba zbudowania K sieci i ich trenowania, w których każda odpowiada za rozpoznanie jednej klasy. Istnieją jednak metody, które wieloklasowość mają wbudowaną w algorytm działania.

Niezależnie od użytego jądra, doboru parametrów sieci, wyboru strategii uczenia, główny problem obliczeniowy w sieciach SVM sprowadza się do rozwiązania zadania programowania kwadratowego z ograniczeniami liniowymi. Zadanie to jest dość dobrze opracowane i istnieje duża liczba klasycznych metod prowadzących do osiągnięcia minimum globalnego funkcji celu. Problem pojawia się przy dużej liczbie danych uczących, gdzie klasyczne algorytmy mają problem z wykorzystaniem pamięci i złożonością obliczeniową. Wymagane są wtedy bardziej wyrafinowane algorytmy przygotowane specjalnie pod kątem działania sieci SVM, do których możemy zaliczyć algorytmy SMO (ang. *Sequential Minimal Optimization*), należące do grupy programowania sekwencyjnego.

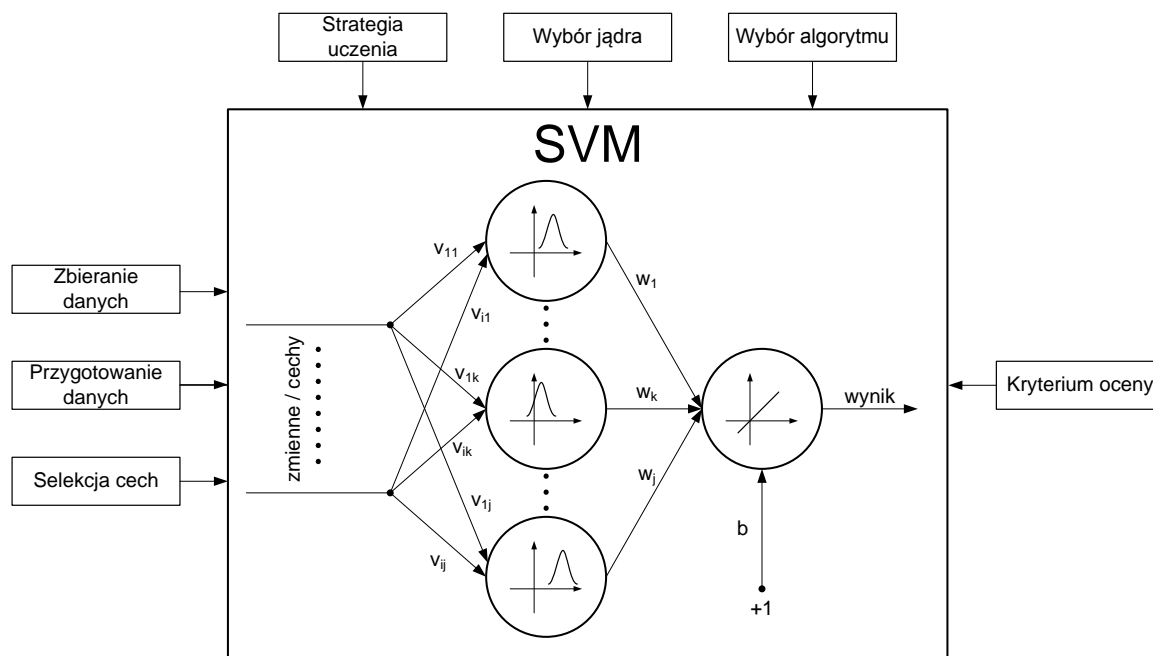
Należy również zaznaczyć, że ogólnie wszystkie algorytmy z reguły są zbieżne do minimum globalnego funkcji celu, ale źle dobrane parametry sieci, które są stałe podczas procesu uczenia, mogą prowadzić do rozwiązania dalekiego od optymalnego. Aby uniezależnić się od tego typu problemu powstało szereg algorytmów, które pozwalają na dobór parametrów sieci w procesie uczenia na podstawie określonych kryteriów założonych na wstępie, przykładem może być dość popularny *nu*-SVM algorytm.

Wynik działania sieci SVM ocenia się za pomocą kryteriów oceny jakości (ang. *performance indices*) innych dla klasyfikatorów i innych dla predyktorów. Zadaniem tych miar jakości, jest pomoc w znalezieniu najlepszego dopasowania modelu do danych uczących przy jak najmniejszej złożoności sieci, co z reguły prowadzi do dobrych właściwości generalizacyjnych.

Problemem pojawiającym się zawsze podczas predykcji i klasyfikacji, którego również nie można zbagatelizować, jest zbieranie oraz przygotowanie danych (ang. *data collection and preprocessing*), które także znacząco wpływa na proces uczenia i testowania modelu. Na przykład w układach elektrycznych ważne jest czy dane użyte w modelowaniu zostały zebrane z eksploatacyjnej pracy urządzenia czy są wynikiem sztucznego wymuszania sygnałów na wejściu. Konieczna jest zwykle normalizacja danych, powodująca znaczną poprawę działania sieci. Ważna i stosowana jest również

standaryzacja danych, określenie przypadków odstających, analiza statystyczna, jak również ocena jakości danych otrzymanych z pomiarów.

W zależności od postawionego celu, rozwiązanie powyżej opisanych problemów pozwala na otrzymanie dobrych rezultatów predykcji lub klasyfikacji przy użyciu sieci SVM. Ogólny schemat przedstawiający problemy, które znacząco wpływają na wynik działania sieci SVM w zadaniach predykcji i klasyfikacji pokazano na rysunku 1.



Rysunek 1. Schemat przedstawiający ogólne problemy związane z wykorzystaniem sieci SVM

4.2 CEL PODJĘTYCH BADAŃ

Celem badań przedstawionego jako cykl publikacji było udowodnienie, że przy prawidłowym przygotowaniu danych, rozwiązaniu problemów doboru optymalnego wektora cech, oraz właściwym doborze struktury i strategii uczenia w połączeniu z wiedzą dotyczącą elektrotechniki i bioinżynierii, sieć SVM jest skutecznym narzędziem, znacznie poprawiającym dotychczasowe wyniki osiągnięte w zadaniach predykcji i klasyfikacji przy zastosowaniu innych metod modelowania.

4.3 WYNIKI BADAŃ I DYSKUSJA

Już podczas studiów magisterskich moje zainteresowania skupiały się wokół sztucznej inteligencji a w szczególności sieci neuronowych, które z czasem ewoluowały w stronę nowych technik jakimi są sieci SVM. Widząc olbrzymi potencjał jaki posiadają, postanowiłem wykorzystać je do rozwiązania problemów, które tradycyjnymi technikami nie zawsze dało się skutecznie rozwiązać. W badaniach skupiłem się na zastosowaniach sieci SVM w dwóch dziedzinach: elektrotechnice i bioinżynierii. Należy zauważyć, że zadania bioinżynierii rozważane w moich pracach opierają się również na sygnałach czujników przetwarzających wielkości nieelektryczne w elektryczne.

Pierwszą pracą poddaną ocenie jest publikacja [1], w której zaproponowałem udoskonaloną metodę lokalizacji zwarć w liniach energetycznych przy wykorzystaniu sieci SVM. Najczęściej wykorzystywane metody lokalizacji zwarć, tzw. metody impedancyjne opierają się na obliczeniach pozwarciowych prądów i napięć fazowych dla podstawowej harmonicznej. Wadą tych rozwiązań jest to, że podstawowe harmoniczna prądu i napięcia są silnie zależne od impedancji zwarcia, która jest w ogólności nieznana. Rezultaty działania algorytmów na tak opartych założeniach powodują, że ich dokładności lokalizacyjne są rzędu kilometrów, co wydłuża czas potrzebny na znalezienie miejsca zwarcia dla celów inspekcyjno-remontowych. Zaproponowałem więc metodę hybrydową do poprawienia dokładności lokalizacji zwarć, składającą się z sieci SVM oraz informacji o wysokich częstotliwościach, które są często pomijane przez badaczy. Analiza stanów nieustalonych występujących po zwarcu pokazała, że pojawiające się wysokie harmoniczne w przebiegach napięć są zależne w dużym stopniu od miejsca zwarcia. Aby je wyodrębnić, zaproponowałem wykorzystanie filtra górnoprzepustowego eliptycznego rzędu ósmego, który po filtracji uwypuklił zakres częstotliwości, w którym występuje maksimum globalne. Znaleziona częstotliwość, dla której występuje maksimum okazała się funkcją miejsca zwarcia i wysoce niezależną od impedancji zwarcia. W wyniku symulacji wystąpienia zwarć w różnych miejscach przy użyciu programu ATP-EMTP uzyskałem charakterystykę częstotliwościową dla całej długości linii przedstawiającą rozkład częstotliwości maksimum w funkcji miejsca zwarcia. Powstała w ten sposób charakterystyka ma przebieg niemonotoniczny i nie mogła być bezpośrednio użyta w procesie uczenia sieci SVM. Cecha ta, jak również amplituda prądu i napięcia dla podstawowej harmonicznej faz „chorych” oraz czas wystąpienia zwarcia zmierzone tylko na początku linii, zostały wykorzystane do zbudowania wektora cech dla sieci SVM w metodzie hybrydowej. Tak zbudowany algorytm lokalizacji zwarć przebadłem na linii energetycznej 400 kV o długości 200 km dla wszystkich możliwych typów zwarć mogących się pojawić na całej długości linii, jak również różnej kombinacji połączenia systemów. Wyniki uzyskane przy użyciu mojej metody wykazały, że średnie błędy bezwzględne lokalizacji zwarcia dla dowolnego typu są poniżej 100 m, a w ujęciu ogólnym dla każdego połączenia systemu poniżej 50 m. Porównanie go z algorytmem Takagi, który jest często cytowany w literaturze wykazało wyższość metody hybrydowej średnio o rząd wielkości. Warty podkreślenia jest fakt, że dla tak przygotowanego wektora cech, sieć SVM w trybie testowym tylko 153 razy pokazała błąd wyższy niż 1 km dla ponad 80000 przypadków testowych. Według mojej wiedzy było to pierwsze takie zastosowanie sieci SVM na świecie do lokalizacji zwarć, co potwierdza duża liczba cytowań dla tego artykułu (wg. Google scholar ponad 100 cytowań od 2004).

Nowe podejście do lokalizacji uszkodzeń w filtrach analogowych zaproponowałem w artykule [2]. Opierając się jedynie na informacji o sygnałach z zacisków zewnętrznych, przeprowadziłem identyfikację parametryczną uszkodzonych elementów przy wykorzystaniu sieci SVM. Poprawność metody przetestowałem na bi-kwadratowym filtrze analogowym RC oraz filtrze drabinkowym RLC dziewiątego rzędu, w których uszkodzenie elementu było skojarzone z tolerancją pozostałych

elementów przyjętą w zakresie do $\pm 5\%$. Identyfikacja parametryczna (tzw. nieidealne uszkodzenie) w tym przypadku jest zadaniem bardzo trudnym, gdyż przyjęcie tolerancji wprowadza rozmycie charakterystyk częstotliwościowych, co utrudnia wykrycie elementu uszkodzonego jak i stwierdzenie stanu normalnego działania układu. Bezpośrednie zastosowanie zmierzonych sygnałów wejściowego i wyjściowego do wykrycia uszkodzenia, okazało się mało skuteczne ze względu na duże rozbieżności sygnałów pomierzonych. Zatem wybór cech dla sieci SVM był zadaniem trudnym, gdyż potraktowanie filtrów analogowych jako czarnej skrzynki daje nam dostęp jedynie do dwóch sygnałów: prądu wejściowego i napięcia wyjściowego. Zaproponowałem więc wykorzystanie charakterystyk częstotliwościowych jako sygnałów wejściowych dla sieci SVM. Wybór częstotliwości, które były podane na wejście sieci SVM przeprowadziłem wykorzystując analizę wrażliwości napięcia wyjściowego na zmiany wartości poszczególnych parametrów. Uwzględniając tylko te częstotliwości, dla których analiza wrażliwości dała maksimum dla wartości dodatnich bądź minimum dla wartości ujemnych, i przyjmując jako cechę amplitudę i fazę, zarówno prądu jak i napięcia otrzymałem wektor cech podawany na wejście sieci SVM. Aby sprawdzić, czy wektor cech jest optymalny, wykorzystałem analizę PCA do obserwacji graficznej rozkładu stanów uszkodzenia przy różnym stopniu odkształcenia wartości parametrów elementów filtru. Analiza PCA wykazała, że najlepszą separację różnych uszkodzeń dla filtru RC otrzymano dla wektora cech składającego się z modułu napięcia wyjściowego, fazy napięcia wyjściowego i modułu prądu wejściowego, a dla filtru RLC – modułu napięcia wyjściowego i modułu prądu wejściowego. Dla tak przygotowanych wektorów cech sieć SVM osiągnęła bardzo dobre wyniki identyfikacji uszkodzeń elementów w całym zakresie ich zmian parametrycznych. Wyniki na poziomie 2.37% dla 1% tolerancji nieuszkodzonych elementów oraz 1.7% dla 5% tolerancji dla filtru RC i 0.71% dla 5% tolerancji dla filtru RLC potwierdziły bardzo dużą skuteczność sieci SVM, znacznie lepszą niż inne metody w podobnych zadaniach. Dodatkową przewagą wykorzystania sieci SVM były jej dobre właściwości generalizacyjne, jak również szybkość wykrywania uszkodzenia w trybie testowym, co pozwalałoby na jej ewentualne zastosowanie w aplikacjach przemysłowych.

Problem, którego rozwiązanie zaproponowałem w artykule [3] dotyczył wyznaczenia optymalnej struktury geometrycznej silnika reluktancyjnego przełączalnego z toczącym się wirnikiem (SRPTW) za pomocą metody elementów skończonych (MES). Wyznaczenie optymalnej struktury geometrycznej silnika za pomocą MES jest zadaniem czasochłonnym i kosztownym ze względu na duże potrzeby obliczeniowe. Zaproponowałem więc użycie sieci SVM do wspomagania obliczeń optymalnych parametrów ruchowych, redukując dzięki temu czas potrzebny na otrzymanie wyników, jak również umożliwiając przeprowadzenie dokładniejszych symulacji. Pierwszym etapem zaprojektowania napędu było dokonanie obliczeń rozkładu stacjonarnego pola magnetycznego 2D o przekroju prostokątnym SRPTW do osi obrotu jego wirnika przy wykorzystaniu MES. Następnie dokonano obliczeń momentów reluktancyjnego (T_r) i mimośrodowego elektromagnetycznego (T_e) dla wybranych zmian wartości konstrukcyjnych parametrów, tj. szerokości zęba wzbudnika (X_{sz}) i

szerokości żłobka wzbudnika (a). Powstałe w ten sposób zależności $Tr=f(Xsz,a)$ i $Te=f(Xsz,a)$ nie dały jednoznacznej odpowiedzi dla jakich wartości parametrów momenty osiągną wartość maksymalną. Zadaniem sieci SVM było znalezienie możliwie zawężonego zakresu parametrów Xsz oraz a , dla których występuje maksymalny moment Tr i Te . Uczenia sieci SVM dokonano w oparciu o wyniki analizy MES, gdzie przygotowany wektor wejściowy składał się z Xsz oraz a , natomiast sygnałem wyjściowym sieci był moment Tr lub Te . Dla tak nauczonych sieci SVM, błąd testowania dla momentu Te wyniósł 6.3% a dla Tr 7.1%. W oparciu o zbudowany model przeprowadzono ponowną symulację o bardzo dużym zagęszczeniu dla parametrów Xsz oraz a . Wyniki pokazały, że należy przeprowadzić ponowne obliczenia momentów zwiększając liczbę danych symulacji w zawężonym zakresie wskazanym przez sieć SVM przy wykorzystaniu analizy MES. Wykorzystanie sieci SVM do rozwiązania problemu związanego z dokładnością i czasem modelowania bardzo złożonych zjawisk fizycznych potwierdziło ich skuteczność jako narzędzia, które idealnie nadaje się do wspomaganie innych bardziej złożonych metod obliczeniowych.

Następnym problemem, którego rozwiązanie zaproponowałem w pracach [4,5] było wykorzystanie sieci SVM do zamodelowania działania układu proporcjonalno-całkująco-różniczkującego (PID) zaimplementowanego w programowalnych sterownikach logicznych (PLC) o nie do końca znanej strukturze. Brak lub niepełna wiedza na temat struktury algorytmu PID powoduje, że pojawiają się problemy z jego dostrojeniem. Według O'Dweyer'a około 50 niestandardowych struktur PID jest implementowanych w sterownikach PLC a producenci często nie do końca informują o dokładnej ich transmitancji, gdyż uważają to za tajemnicę handlową. Często układy te są również zmodyfikowane i zawierają takie udogodnienia jak anti-windup, filtrowanie wartości zadanej, auto-tuning, wprowadzenie strefy nieczułości itp. Dla tak zmodyfikowanych rozwiązań PID brak propozycji nastaw, które z reguły są podawane tylko dla klasycznych struktur. Aby rozwiązać powyższy problem, zaproponowałem użycie sieci SVM, której zadaniem było dokładne modelowanie odpowiedzi dowolnego algorytmu PID zaimplementowanego w sterowniku PLC dla dowolnych nastaw. Rozwiązanie takie dawałoby możliwość wykorzystania na przykład algorytmów genetycznych do doboru optymalnych nastaw dla tak zamodelowanego układu.

Pierwsze próby zamodelowania działania układu PID przedstawiłem w pracy [4]. Zaproponowany w niej wektor cech spowodował, że sieć SVM dobrze symulowała działanie układu PI. Rozszerzenie go do ogólnej struktury PID nie dało jednak tak dobrych wyników i w pracy [5] zaproponowałem nowe podejście do selekcji cech oparte o model nieliniowej autoregresji z sygnałem zewnętrznym (NARX). Pierwszym krokiem nowego podejścia było przygotowanie danych uczących i testujących. Rozwiązanie polegało na rezygnacji z pozyskania danych z normalnej pracy algorytmu PID w pętli sprzężenia zwrotnego na rzecz sztucznego wymuszania sygnałów pojawiających się na wejściu algorytmu PID. W modelu NARX głównym problemem jest dobór regresorów, czyli liczby próbek „wstecz” sygnałów wejściowych i wyjściowych. Nowatorskim podejściem jakie zaproponowałem w rozwiązaniu był sposób wyboru optymalnej liczby regresorów, który opierał się na zbudowaniu

dwuwymiarowej mapy rozkładu błędu MSE dla wszystkich kombinacji regresorów. Na podstawie tak zbudowanych map, wybrałem optymalną liczbę regresorów, a wyniki sieci SVM jakie otrzymałem dla tej metody okazały się lepsze niż dla modeli transmitancji dostarczonych przez producentów. Poprawność metody przetestowałem na dwóch różnych sterownikach PLC i trzech algorytmach PID w nich zaimplementowanych. Metoda ta okazała się na tyle uniwersalna, że może być zastosowana do strojenia różnych algorytmów PID z dowolnymi modyfikacjami wprowadzonymi przez producenta.

Rozwiązanie dwóch istotnych problemów z zakresu bioinżynierii dotyczących optymalizacji doboru dawek leków podawanych organizmom żywym przedstawiłem w cyklu trzech artykułów [6,7,8], proponując autorską metodę modelowania silnie nieliniowej zależności dawka-efekt na podstawie sygnałów elektrycznych w tak zwanym teście maksymalnego wstrząsu elektrycznego (MES). Wykorzystanie sieci SVM jest w tym wypadku bardzo wskazane, gdyż ze względów przede wszystkim etycznych niedozwolone jest testowanie w nieskończoność dawek eksperymentalnych i obserwowania, jaki efekt wywierają one na organizm zwierząt.

Pierwszy problem dotyczył wyznaczania progów toksyczności dla leków przeciwpadaczkowych. Jego rozwiązanie zaproponowałem w pracy [6]. Każdy lek po przekroczeniu pewnego bezpiecznego przedziału dawkowania zaczyna wywierać efekt toksyczny. Zauważono, że leki przeciwpadaczkowe w zakresie wysokich dawek mogą wywierać paradoksalnie działanie prodrgawkowe, a więc zamiast zmniejszać liczbę napadów drgawek czyli podnosić próg pobudliwości drgawkowej układu nerwowego, mogą obniżyć go, co sprzyja pojawianiu się drgawek. Badania przeprowadziłem na pięciu związkach o aktywności przeciwdrgawkowej wykorzystując model drgawek indukowanych prądem elektrycznym tzw. MES test. Dane zebrane do badań pochodziły z obserwacji zwierząt, którym podawano różne dawki leku i wywoływano drgawki stałym prądem elektrycznym z zakresu 24-60 mA licząc wyprosty kończyn tylnych. Z punktu widzenia medycznego istotne jest znalezienie takiej dawki leku, która hamuje pojawianie się wyprostów kończyn u 50% grupy zwierząt dla danego natężenia prądu. Zaproponowałem wykorzystanie sieci SVM do zamodelowania zależności (dawka, liczba wyprostów – natężenie prądu), co było trudnym zadaniem ze względu na silną nieliniowość takiej zależności i wymagało starannego przygotowania danych uczących. Dodatkową trudnością była bardzo ograniczona liczba danych pomiarowych. Ostatecznie po przebadaniu kilku struktur sieci SVM, wyniki błędu otrzymane w uczeniu nie przekraczały 3.5%, a w testowaniu 8.2%. Sieć nauczona do aproksymacji powyższej zależności została następnie wykorzystana do predykcji natężenia prądu CS_{50} (ang. *current strength*), jaki indukuje napady drgawkowe u 50% badanej grupy przy określonej dawce leku. Na podstawie tak przeprowadzonej predykcji można wyznaczyć tzw. okno terapeutyczne, czyli przedział między dawką toksyczną a terapeutyczną, co było jednym z postawionych celów. Wyznaczone teoretycznie natężenie prądu CS_{50} pozwoliło na obliczenie optymalnej dawki dla każdego z badanych leków, stanowiącej próg ich toksyczności. Ostatecznie obliczone dawki zostały podane zwierzętom eksperymentalnym. Otrzymane wyniki różniły się o co najwyżej 3%, co jednoznacznie potwierdziło skuteczność takiego wykorzystania sieci SVM.

Zachęcony dobrymi wynikami działania sieci SVM zaproponowałem rozszerzenie metody na wyznaczanie progu wrażliwości bólowej dla interakcji lekowej, która została przedstawiona w pozycji [7]. Kombinacje leków (interakcje leków) są często wykorzystywane w medycynie do osiągnięcia zamierzonego efektu bez konieczności używania wysokiej dawki pojedynczego leku. W porównaniu do leków stosowanych w monoterapii, politerapia często umożliwia optymalizację farmakoterapii - to jest uzyskanie efektu leczniczego poprzez zastosowanie niższych dawek dwóch leków podanych łącznie. Problemem jaki należy tu rozwiązać jest dobór dawek dla kombinacji w taki sposób, aby otrzymać efekt zbliżony do efektu pojedynczo podanej dawki. Ponieważ badania dotyczą organizmów żywych, nie ma możliwości zbudowania modelu matematycznego opisującego w ujęciu ogólnym wszystkie reakcje zachodzące w organizmie. Zaproponowałem więc zbudowanie modelu zależności dawka – efekt dla różnych kombinacji leków. Dane potrzebne do uczenia pochodziły z badań *in vivo* dla dwóch związków podawanych pojedynczo i w kombinacji. Wynikiem podawania tych związków było wyznaczenie progu wrażliwości bólowej mierzonej za pomocą sygnałów elektrycznych z mostka tensometrycznego (tzw. test von Frey'a). Otrzymane dane po normalizacji, odrzuceniu statystycznie nieistotnych danych i błędów pomiarowych, posłużyły do uczenia i testowania sieci SVM. Ponieważ danych do uczenia było niewiele zastosowałem 10-krotną walidację krzyżową umożliwiającą obiektywną ocenę zdolności generalizacji sieci. W efekcie otrzymałem model SVM, dla którego błąd względny testowania osiągnął wartość 7.17%. Dla tak otrzymanego modelu przeprowadziłem predykcję kombinacji dawek w całym zakresie, gdzie suma podania dawek nie przekraczała maksymalnej dawki pojedynczego związku (30 mg/kg m.c.). Otrzymane w ten sposób wyniki z sieci SVM pozwoliły na wyznaczenie dawek, dla których wystąpiła maksymalizacja efektu podania, tj. najwyższy próg wrażliwości bólowej. Wyniosły one odpowiednio 4 mg/kg m.c. i 1 mg/kg m.c. Efekt jaki wywołało podanie tej kombinacji był porównywalny z efektem podania 30 mg/kg m.c. pojedynczego związku. Wyniki otrzymane przy użyciu SVM sprawdzono w warunkach *in vivo*, potwierdzając ostatecznie poprawność zaproponowanej metody.

Dobre właściwości modelu dawka–efekt opartego na sieci SVM pozwoliły na nowatorskie wykorzystanie ich w analizie izobolograficznej, które wyniki przedstawiłem w artykule [8]. Analiza izobolograficzna służy do ilustracji graficznej zachowania się kombinacji leków. Na jej podstawie możemy stwierdzić, czy działanie kombinacji leków ma charakter addytywny, super-addytywny czy sub-addytywny, co ma duże znaczenie w farmakologii i farmakoterapii. Wykonanie analizy metodą klasyczną wiąże się z znaczną liczbą przeprowadzonych eksperymentów i dużą liczbą wykorzystanych zwierząt, co jest sprzeczne z obowiązującymi wytycznymi komisji etycznych ds. badań na zwierzętach. Aby ograniczyć czas i koszty wykonania analizy, badania często ograniczane są do jednej, bądź dwóch kombinacji leków. Wykonanie tak małej liczby kombinacji nie daje wglądu w naturę interakcji w ujęciu ogólnym.

Zbudowany za pomocą sieci SVM model (dawka–efekt) pozwolił na symulację różnych proporcji dawek bez konieczności wykonywania dalszych eksperymentów *in vivo*. Badania przeprowadzono dla

dwóch związków, z których jeden nie miał jeszcze znanego mechanizmu działania. Zaletą wykorzystania sieci SVM było szybkie oszacowanie kombinacji dawek, które były odpowiednikiem dawek podawanych oddzielnie. Zbudowany izobologram przy wykorzystaniu sieci SVM dał wgląd w ogólną naturę interakcji, jak również pozwolił na pośrednie ustalenie mechanizmu działania nowej substancji chemicznej, która nie miała zdefiniowanego mechanizmu działania. Poprawność metody potwierdzono w eksperymentach *in vivo*. Metoda ta okazała się na tyle uniwersalna, że może być stosowana dla różnych poziomów efektów i według mojej wiedzy było to nowatorskie użycie sieci SVM, po raz pierwszy zaproponowane w tej dziedzinie badań.

Ostatni przykład wykorzystania sieci SVM przedstawiłem w pracy [9] gdzie zaproponowałem inteligentną metodę doboru optymalnych ustawień procesu zgrzewania przy wykorzystaniu elektrycznej zgrzewarki tarciovej. Problem jaki został rozwiązany, dotyczył doboru takich parametrów elektrycznej zgrzewarki, aby podczas zgrzewania dwóch próbek żeliwa otrzymać połączenie o jak największej wytrzymałości na rozciąganie. Sam proces zgrzewania żeliwa jest procesem złożonym i brakuje dokładnego modelu fizycznego dobrze go opisującego. Według Amerykańskiego Towarzystwa Spawalniczego połączenie żeliwa za pomocą zgrzewania tarciowego jest bardzo trudne do osiągnięcia, ponieważ grafit znajdujący się w żeliwie zachowuje się jak smar i zapobiega wytworzeniu się odpowiedniej ilości ciepła potrzebnego do połączenia. Do rozwiązania tego problemu zaproponowałem użycie sieci SVM, której zadaniem było zamodelowanie zjawiska zgrzewania na podstawie jedynie sygnałów wejściowych będących (siłą, czasem zgrzewania i siłą spęczania) oraz wyjściowego (wytrzymałość na rozciąganie). Za dane służące do modelowania posłużyły wyniki eksperymentalne procesu zgrzewania przeprowadzonego przez eksperta. Przeprowadzono czterdzieści dwa eksperymenty, w wyniku których dla różnorodnych nastaw otrzymano różne wartości wytrzymałości na rozciąganie. Najlepszy wynik jaki otrzymano wyniósł 211 MPa. Tak otrzymane dane eksperymentalne użyto do zbudowania modelu opartego o sieć SVM, dla którego błąd względny dla danych testujących został zredukowany do wartości 2.6%. Zbudowany model sieci SVM o dobrych właściwościach generalizacyjnych, posłużył następnie do definicji funkcji celu dla algorytmów genetycznych. Optymalizacja nastaw parametrów wejściowych za pomocą algorytmów genetycznych na modelu powstałym na sieci SVM pozwoliła uzyskać o 21.5% lepszą wytrzymałość na rozciąganie równą 256.93 MPa w stosunku do wyników uzyskanych przez eksperta. Poprawność zaproponowanej metody potwierdziły również wyniki eksperymentalne, gdzie za nastawy elektrycznej zgrzewarki tarciovej przyjęto parametry obliczone przy użyciu algorytmów genetycznych. Wynik wytrzymałości na rozciąganie równy 258 MPa różnił się jedynie o mniej niż 1% od teoretycznie wyliczonego przez metodę połączenia sieci SVM i algorytmów genetycznych.

4.4 Podsumowanie

Artykuły tworzące cykl publikacji powiązanych tematycznie oceniło w sumie 20 recenzentów z całego świata. Cechą wspólną łączącą wszystkie zaprezentowane tu rozwiązania jest zastosowanie

podobnych narzędzi i technik sztucznej inteligencji opartych na zastosowaniu maszyny wektorów nośnych SVM.

Na podstawie zaproponowanych metod, opracowaniu własnej metodyki badawczej i głębokiego zrozumienia problemów udało mi się rozwiązać szereg trudnych zadań z zakresu elektrotechniki i bioinżynierii. Do najważniejszych osiągnięć zaliczyłbym:

- Opracowanie i późniejsze udoskonalenie dokładnej metody lokalizacji zwarć w liniach energetycznych, która w czasach publikacji pracy przewyższała znane techniki impedancyjne o rząd wielkości. Według mojej wiedzy było to pierwsze takie opisane w literaturze zastosowanie sieci SVM do lokalizacji zwarć.
- Opracowanie nowego podejścia do lokalizacji uszkodzeń w filtrach analogowych, dzięki któremu można szybko i z dużą dokładnością lokalizować uszkodzone elementy o różnym stopniu uszkodzenia (tzw. uszkodzenia parametryczne).
- Opracowanie metody wspomaganego modelowania parametrów konstrukcyjnych w maszynach elektrycznych, co przyczyniło się do zwiększenia dokładności i przyspieszenia procesu modelowania złożonych zjawisk fizycznych.
- Opracowanie nowatorskiego podejścia do modelowania struktur PID o nieznanych transmitancjach operatorowych zaimplementowanych w sterownikach PLC, ułatwiającego dostrajanie parametrów układu.
- Opracowanie metod wykorzystujących sieci SVM do modelowania progów toksyczności leków, jak również doboru kombinacji leków w przyjętych modelach bólu. Wykorzystanie jednej z tych metod w izobolografii przyczyniło się istotnie do obniżenia koniecznej liczby przeprowadzonych eksperymentów na zwierzętach.
- Opracowanie metody doboru optymalnych parametrów procesu zgrzewania tarcowego w zgrzewarce elektrycznej przy wykorzystaniu sieci SVM i algorytmu genetycznego.

Opracowane przeze mnie zaawansowane metody przetwarzania sygnałów przy użyciu metod sztucznej inteligencji, wydają się na tyle uniwersalne, że z powodzeniem mogą znaleźć zastosowanie w innych dziedzinach badań, np. medycynie, ekonomii, zarządzaniu itp.

