

Raport Badawczy
Research Report

RB/60/2013

**Podsystem handlu energią
elektryczną z operatorem sieci
zewnętrznej
Metoda generowania cen
przez operatora**

J. Stańczak

Instytut Badań Systemowych
Polska Akademia Nauk

Systems Research Institute
Polish Academy of Sciences



POLSKA AKADEMIA NAUK

Instytut Badań Systemowych

ul. Newelska 6

01-447 Warszawa

tel.: (+48) (22) 3810100

fax: (+48) (22) 3810105

Kierownik Zakładu zgłaszający pracę:
Prof. dr hab. inż. Zbigniew Nahorski

Warszawa 2013

INSTYTUT BADAŃ SYSTEMOWYCH
POLSKA AKADEMIA NAUK

Jarosław Stańczak

Podsystem handlu energią elektryczną z operatorem sieci zewnętrznej
Metoda generowania cen przez operatora

Warszawa 2013

Projekt badawczy własny Nr N N519 580238
Ministerstwa Nauki i Szkolnictwa Wyższego

Komputrowe zarządzanie energią w ośrodku badawczo-szkoleniowym z rozproszonymi źródłami energii i zmiennym zapotrzebowaniem energetycznym na eksperymenty badawcze

Kierownik projektu:

prof. dr hab. inż. Zbigniew Nahorski

Zadanie:

Opracowanie algorytmów do handlu na bieżąco energią elektryczną z operatorem sieci rozdzielczej średniego napięcia

1. Wstęp

Projekt systemu zarządzającego przepływem energii elektrycznej w ośrodku badawczo-szkoleniowym zawiera moduł odpowiedzialny za handel z operatorem sieci zewnętrznej niedoborami lub nadwyżkami wyprodukowanej energii. Moduł ten ma pełnić dwojakie funkcje. W horyzoncie długoterminowym powinien stanowić element doradczy dla operatora, zawierającego kontrakty długoterminowe. W horyzoncie krótszym (do doby) ma samodzielnie zawierać kontrakty i być zintegrowany w systemie agentowym, tworzącym wraz z innymi wytwórcami i odbiorcami rynek energii dla niedużych producentów i odbiorców, szczególnie czystej energii ze źródeł odnawialnych.

Niestety, jednym z głównych problemów, jakie napotyka się przy realizacji tego elementu systemu, jest brak uregulowań prawnych, a co za tym idzie brak takiego rynku w warunkach polskich. Istniejące w chwili obecnej giełdy i rynki energii są stworzone dla wielkich producentów i tak mały wytwórca jak centrum badawczo-szkoleniowe z przyczyn formalnych nie mogłoby brać w nich udziału. Wobec tego w pracy niniejszej zaproponowano rozwiązanie, które umożliwi zaprojektowanie uruchomienie i zasymulowanie takiego systemu nawet w sytuacji braku odpowiedniego rynku w rzeczywistości. Siłą rzeczy pewne założenia przyjęte w tym rozwiązaniu mogą się nie sprawdzić lub znacznie odbiegać od rzeczywistości, gdy rynek taki powstanie.

2. Założenia

W związku z brakiem odpowiedniego rynku dla małych wytwórców energii elektrycznej, do stworzenia odpowiedniego modelu posłużono się wzorem istniejącego rynku energii dla jej dużych producentów. Jako rynek wzorcowy wybrany został rynek dnia następnego, istniejący od 30 czerwca 2000, którego dane z przeprowadzonych transakcji uśrednione dla okresów godzinnych są dostępne w Internecie (<http://wyniki.tge.pl/pl>). Na podstawie zamieszczonych tam cen przeprowadzonych transakcji zbudowany został moduł proponujący ceny dla systemu agentowego, operującego na modelowanym rynku.

Jako moduł doradczy wybrana została sieć neuronowa typu BP. Sieć ma za zadanie nauczyć się danych o przeprowadzonych transakcjach, a nauczywszy się ich wykrywać wszelkie regularności, okresowości i trendy w nich występujące, tak aby stać się „ekspertem” od wyceny jednostek sprzedawanej energii.

3. Zastosowana sieć neuronowa

Do nauki szeregów czasowych, reprezentujących ceny zawartych transakcji na dostawę prądu, wybrana została sieć neuronowa typu backpropagation (BP). Jest to dość typowe zastosowanie tego typu sieci, gdyż bardzo dobrze radzi sobie ona z nauczeniem się danych tworzących szeregi czasowe, a potem z wykorzystaniem nauczonych informacji do przetwarzania sygnałów wejściowych, których nie było wśród danych trenujących. Dzieje się tak dlatego, że sieć typu BP posiada zdolność generalizacji, czyli potrafi znaleźć (najczęściej dobre) odpowiedzi dla niewyuczonych sekwencji danych na wejściu.

3.1 Budowa i działanie sieci typu BP

Istnieje wiele odmian sieci neuronowych, różniących się sposobem uczenia, kierunkiem przepływu sygnału i metodami uczenia. Wykorzystywana w pracy sieć typu BP należy do kategorii sieci jednokierunkowych (bez sprzężeń zwrotnych), uczonych z nauczycielem przy

użyciu algorytmu znanego pod nazwą „backpropagation”, od którego sieć bierze swoją nazwę.

Sieć typu BP składa się z elementów zwanych sztucznymi neuronami. Są to jednostki przetwarzające sygnały. Są one wzorowane na komórkach nerwowych występujących na przykład w ludzkim mózgu. Odpowiednio połączone ze sobą neurony tworzą sieć neuronową, która posiada zdolność nauki oraz przetwarzania danych na podstawie wyuczonych wzorców.

Pojedynczy sztuczny neuron (zwany dalej dla uproszczenia neuronem) składa się z szeregu wejść, na które otrzymuje sygnały wyjściowe z innych neuronów bądź z zewnątrz, zbioru wag, czyli pewnych współczynników przypisanych do konkretnych wejść, przez które przemnażana jest wartość sygnału wejściowego (jedno dodatkowe wejście, również zaopatrzone w wagę, jest na stałe podłączone do sygnału o wartości 1). Kolejnym elementem neuronu jest sumator ważonych sygnałów wejściowych. Po zsumowaniu sygnał przetwarzany jest przez tak zwaną funkcję aktywacji lub przejścia neuronu. Najczęściej wykorzystywaną funkcją aktywacji jest funkcja sigmoidalna o wzorze

$$y(x) = 1 / (1 + \exp(-\alpha x)) \quad (1),$$

która charakteryzuje się nasyceniem sygnału zarówno dla wejść bardzo małych, jak i bardzo dużych oraz pewnym odcinkiem, dla którego następuje płynne przejście między wartościami nasycenia. Stosuje się także inne funkcje aktywacji neuronu, w tym także progowe i liniowe. Tak uformowany sygnał staje się wyjściem neuronu. Można to zapisać w postaci wzoru [2, 3]:

$$V_{ij}^{\mu} = y \left(\sum_{l=1}^{k_{i-1}} w_{jl} V_{(i-1)l}^{\mu} \right) \quad (2)$$

gdzie:

V_{ij}^{μ} – odpowiedź j -tego neuronu warstwy i na sygnał z warstwy poprzedniej, powstały przy prezentacji wzorca μ ;

i – numer warstwy, przyjęto następującą konwencję numeracji warstw sieci o m warstwach: 1 – warstwa wejściowa, m – warstwa wyjściowa;

$y(\dots)$ – funkcja aktywacji;

w_{jl} – waga między neuronem j warstwy i , a neuronem l warstwy $i-1$ (o liczbie neuronów k_{i-1}).

Pojedynczy neuron nie jest zbyt skomplikowanym tworem, nie jest też specjalnie użyteczny. Siła obliczeń neuronowych tkwi w zestawieniu neuronów w sieć o odpowiedniej strukturze. W przypadku sieci BP neurony są organizowane w warstwy tak, że wejścia warstwy następnej są połączone z wyjściami warstwy poprzedniej. W ogólnym przypadku każdy neuron warstwy następnej otrzymuje sygnał od każdego neuronu warstwy poprzedniej.

Liczba warstw sieci nie jest w żaden sposób określona, jednak zgodnie z twierdzeniem Kołmogorowa ([1, 4]) o aproksymacji funkcji ciągłych, które można zastosować również w przypadku sieci typu perceptronowego, wiadomo, że powinna ona mieć przynajmniej trzy warstwy, aby mogła aproksymować dowolne funkcje ciągłe n zmiennych. Liczność neuronów w każdej z warstw dobierana jest najczęściej doświadczalnie, choć z twierdzenia Kołmogorowa można wywnioskować, że liczba neuronów warstwy ukrytej powinna wynosić $2n$. Jednakże w konkretnych przypadkach możliwe jest znaczne ograniczenie tej liczby i najczęściej dobiera się ją doświadczalnie, na przykład w trakcie uczenia lub przez wielokrotne uczenie sieci o różnych rozmiarach.

Algorytm uczenia sieci typu BP polega na przetwarzaniu pewnego zbioru rozwiązań trenujących, zawierających sygnały wyjściowe do sieci i pożądane w takiej sytuacji wyjścia. Na podstawie różnicy (błędu) między sygnałem wyjściowym a wzorcowym obliczane są współczynniki korekcji wag w kolejnych warstwach, począwszy od wyjściowej. Wagi są zmieniane w taki sposób, aby przy ponownym przedstawieniu takiego wzorca błęd odpowiedzi sieci był mniejszy od otrzymanego poprzednio – celem działania algorytmu

uczącego jest minimalizacja sumy kwadratów błędów popełnianych przez sieć (warstwę wyjściową $i=m$ sieci) ([2, 3]):

$$E(w) = \frac{1}{2} \sum_{\mu=1}^p \sum_{j=1}^n (V_j^\mu - O_j^\mu)^2 \quad (3)$$

gdzie:

$E(w)$ – suma kwadratów błędów wszystkich n wyjść sieci po prezentacji p wzorców;
 V_j^μ – odpowiedź j -tego neuronu warstwy wyjściowej, powstały przy prezentacji wzorca μ ;
 O_j^μ – wzorec μ dla j -tego neuronu warstwy wyjściowej.

Propagowany do warstw wcześniejszych błąd sieci wyraża się następującym wzorem dla warstwy wyjściowej ($i=m$) ([2, 3]):

$$\delta_{mj}^\mu = y_j^\mu \left(\sum_{l=1}^k w_{mjl} V_{(m-1)l}^\mu \right) (V_{mj}^\mu - O_j^\mu) \quad (4)$$

i nieco innym dla warstw ukrytych:

$$\delta_{ij}^\mu = y_j^\mu \left(\sum_{l=1}^{k_{i-1}} w_{ijl} V_{(i-1)l}^\mu \right) \left(\sum_{l=1}^{k_{i+1}} w_{(i+1)jl} \delta_{(i+1)l}^\mu \right) \quad (5),$$

przy czym wagi korygowane są według wzoru:

$$\Delta w_{ijl} = \eta \delta_{(i+1)jl}^\mu V_{(i-1)l}^\mu \quad (6)$$

i

$$w_{ijl}(t+1) = w_{ijl}(t) + \Delta w_{ijl} \quad (7)$$

gdzie:

V_j^μ – odpowiedź j -tego neuronu warstwy i na sygnał z warstwy poprzedniej, powstały przy prezentacji wzorca μ ;
 O_j^μ – wzorec μ dla j -tego neuronu warstwy wyjściowej;
 $y(\dots)$ – funkcja aktywacji;
 w_{ijl} – waga między neuronem j warstwy i , a neuronem l warstwy $i-1/i+1$ (o liczbie neuronów k_{i-1}/k_{i+1});
 δ_{ijl}^μ – propagowany błąd sieci.

Przetwarzanie wzorców przez sieć następuje wielokrotnie. Proces uczenia jest długotrwały i wymaga najczęściej tysięcy powtórzeń, zanim sieć nauczy się klasyfikacji wzorców z akceptowalnie małym błędem. Możliwe jest także nieosiągnięcie zakładanego poziomu błędu, jeśli przyjęto na przykład nieodpowiednią strukturę sieci lub za mało neuronów w warstwach. Algorytm BP uczenia sieci oparty jest na optymalizacji gradientowej funkcji celu, dlatego też jest nieodporny na minima lokalne i jest mało prawdopodobne, aby znalazł optimum globalne. Istnieją różne sposoby polepszenia parametrów poszukiwania wag sieci. Polegają one na zastosowaniu bardziej zaawansowanych algorytmów uczenia, na przykład wykorzystanie optymalizacji metodą gradientów sprzężonych, a nie metody najszybszego spadku, wykorzystanie algorytmów ewolucyjnych do uczenia sieci, czy metoda momentum [3], która była wykorzystywana do obliczeń w niniejszej pracy. Lecz w dalszym ciągu uczenie sieci jest procesem pracochłonnym, o czym można się przekonać patrząc na liczby iteracji występujące na rys. 2.

3.2 Dane wykorzystane do uczenia sieci

W poprzednich rozdziałach wspomniano, że rynek energii dla niewielkich producentów nie powstał jeszcze w naszym kraju. Z tego powodu trudno jest go efektywnie modelować,

nie mając danych opisujących jego parametry lub choćby norm prawnych regulujących jego działanie. Dlatego też zdecydowano się na modelowanie rynku przez analogię do rynków istniejących, a w szczególności rynku dnia następnego. Na podstawie dostępnych w sieci wyników działania rynku można przeanalizować występujące tam trendy, wahania i regularności cen, które zapewne w podobnym stopniu wystąpiłyby też na rynku dla małych producentów. Do przeanalizowania i nauczenia się występujących tam reguł wybrana została sieć neuronowa.

Pobrane do nauki sieci dane zawierają uśrednione ceny kontraktów zawartych na konkretne godziny w kolejnych dniach, tygodniach i miesiącach. Jako podstawę do nauki wybrano paczkę danych zawierających cały ostatni rok działania giełdy, czyli rok 2011. Numer roku nie jest w tych danych wykorzystywany, gdyż sieć uczono danymi z jednego roku. Przetestowano kilka wariantów danych wejściowych sieci z różnymi zestawami parametrów wejściowych:

1. Data, składająca się z czterech liczb całkowitych, podzielona na miesiąc (1-12), dzień (1-31 – zależnie od miesiąca) miesiąca, dzień tygodnia (0-6, czyli niedziela-sobota) i godzinę (0-23) oraz znacznik kategorii dnia: 0 – święto, 1 – sobota (czyli dzień wolny dla większości przedsiębiorstw) i 2 – dzień roboczy.
2. Data, składająca się z trzech liczb całkowitych, podzielona na numer dnia w roku (1-365), dzień tygodnia (0-6, czyli niedziela-sobota) i godzinę (0-23) oraz znacznik kategorii dnia: 0 – święto, 1 – sobota (czyli dzień wolny dla większości przedsiębiorstw) i 2 – dzień roboczy.
3. Data, składająca się z trzech liczb całkowitych, podzielona na numer godziny w roku (1-8760), dzień tygodnia (0-6, czyli niedziela-sobota) i godzinę (0-23) oraz znacznik kategorii dnia: 0 – święto, 1 – sobota (czyli dzień wolny dla większości przedsiębiorstw) i 2 – dzień roboczy.
4. Data, składająca się z dwóch liczb całkowitych, podzielona na numer godziny w roku (1-8760) i dzień tygodnia (0-6, czyli niedziela-sobota) oraz znacznik kategorii dnia: 0 – święto, 1 – sobota (czyli dzień wolny dla większości przedsiębiorstw) i 2 – dzień roboczy.
5. Data, składająca się z trzech liczb całkowitych, podzielona na numer godziny w roku (1-8760), dzień tygodnia (0-6, czyli niedziela-sobota), godzinę (0-23), znacznik kategorii dnia: 0 – święto, 1 – sobota (czyli dzień wolny dla większości przedsiębiorstw) i 2 – dzień roboczy oraz zmienną pomocniczą o wartości 0 dla godzin w roku od 1 do 6999 i wartości 1 dla godzin od 7000 do 8760 (w tym czasie ceny gwałtownie rosną, zapewne z powodu zbliżającego się okresu świąteczno-noworocznego).
6. Wszystkie dane użyte w punktach poprzednich: numer dnia w roku (1-365), miesiąc (1-12), dzień miesiąca (1-31), dzień tygodnia (0-6), numer godziny w roku (1-8760), godzina doby (0-23), kategoria dnia (0-2), zmienna pomocnicza - wskaźnik okresu przedświątecznego (0-1).

Dane wejściowe zostały znormalizowane tak, aby przyjmowały wartości z zakresu (0-1).

Jak można zauważyć, analizując dane uczące, największy wpływ na cenę mają: pora dnia, dzień tygodnia i pora roku. Oczywiście jest to związane ze zmieniającym się zapotrzebowaniem na energię, działalnością firm, koniecznością ogrzania i oświetlenia pomieszczeń, itp.

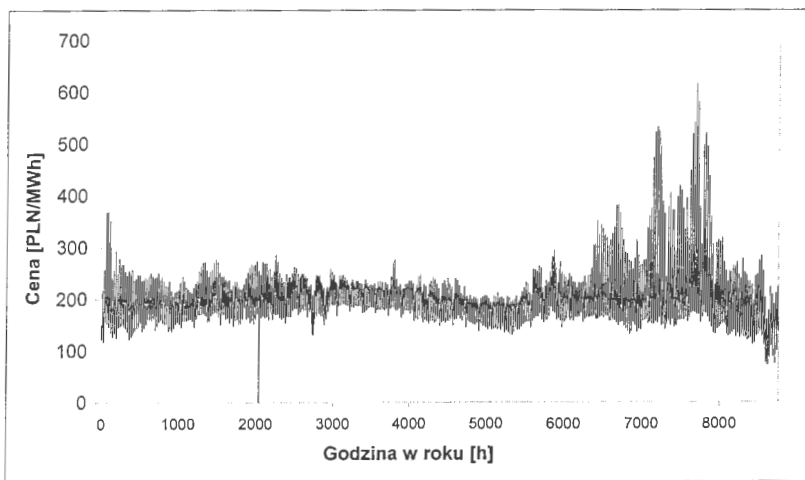
Ponieważ w danych uczących, jak i otrzymywanych wynikach wyraźnie widoczne są różnice cen między dniami świątecznymi i zwykłymi, a jednocześnie dzień tygodnia nie określa jednoznacznie wszystkich dni świątecznych, więc przetestowano również wersję sieci z dodatkowym wejściem, które określa rodzaj dnia. Dla dnia zwykłego sieć otrzymuje na wejściu wartość 2, dla soboty (która jest dla dużej części społeczeństwa dniem wolnym)

podawana jest wartość 1, a dla dni świątecznych i niedziel wartość 0. Pewne problemy może wprowadzać święto Trzech Króli (6.01), obchodzone jako państwowe od roku 2012, a więc niewystępujące w danych uczących (2011). Podobnie można również dyskutować nad wartościami przypisanymi dniom takim jak 2.05 (środek tzw. majowego weekendu), 24.12 i 31.12, itp., które nominalnie są dniami pracy, jednakże wiele przedsiębiorstw działa wtedy w ograniczonym zakresie, co ma z pewnością wpływ na zapotrzebowanie i ceny energii. W danych uczących zastosowano oznaczenie ich jako dni zwykłych, jednakże może to wprowadzać dodatkowe błędy.

Parametrem uczonym jest średnia cena transakcji uzyskana dla danej daty, przedstawiona na Rys. 1. Jest to liczba rzeczywista z zakresu od około 72 do 680. Sporadycznie występują też zera (uwidocznione na Rys. 1), gdy nie doszło do transakcji. Aby można je było wykorzystać do nauki sieci, są one zastępowane przez wartość średnią z sąsiednich wartości, gdyż zera nie są w tym wypadku prawdziwymi danymi (taka cena nie była użyta w żadnej transakcji) i nie powinny być użyte do uczenia sieci. Podobnie konieczne okazało się odfiltrowanie zer w danych testujących, gdyż niepotrzebnie znacznie podbiłyby błąd działania sieci.

Testowano również możliwość uczenia danych znormalizowanych, lecz otrzymywane wyniki nie były dobre.

Zastosowanie sieci do symulacji polega na podawaniu na jej wejście zestawów liczb reprezentujących daty, które nie były uczone, gdyż w kolejnych latach sekwencje liczb na wejściu sieci będą różne. Odpowiedź sieci będzie stanowić daną doradcą, sugerującą wieloagentowemu systemowi handlu cenę transakcji.



Rys. 1. Dane uczące – przebieg ceny energii elektrycznej w ciągu roku 2011.

Analizując powyższe dane widać dość zaskakujący fakt, że najwyższe ceny energia osiąga w listopadzie i na początku grudnia. Ponieważ zjawisko to nie jest do końca uzasadnione przesłankami pogodowymi (równie krótki jest dzień w styczniu i lutym, a temperatury raczej niższe), lecz zwiększonym zapotrzebowaniem na energię w okresie przedświątecznym, więc wprowadzono dodatkową zmienną wejściową (zestaw uczący nr 5 i 6), która ma za zadanie pomóc sieci w wykryciu tej zależności.

3.3 Sieć zastosowana do prognozowania cen energii elektrycznej

Do rozwiązania postawionego problemu została wykorzystana sieć neuronowa typu perceptronu wielowarstwowego o trzech warstwach neuronów: wejściowej, ukrytej i wyjściowej. Sieć zbudowana została z neuronów o nieco innej budowie, zależnie od warstwy. Warstwa wejściowa pełni jedynie rolę dystrybucyjną i nie modyfikuje przechodzących przez nią sygnałów. Warstwa ukryta zbudowana została z neuronów z sigmoidalną funkcją aktywacji. W warstwie wyjściowej zastosowano neurony z liniową funkcją aktywacji, czyli funkcją przekazującą na wyjście ważone sumy sygnałów wejściowych. Użycie liniowej funkcji aktywacji jest związane z wykorzystywanym zakresem wartości wyjściowych, który jest najczęściej zdecydowanie większy niż wartości otrzymywane z funkcji sigmoidalnej. Może więc ona mieć także charakter skalujący sygnały wyjściowe, jednakże w przyjętym rozwiązaniu skalowanie przez tę warstwę nie jest wykorzystywane.

Topologia sieci została dobrana częściowo na podstawie danych uczących, a częściowo doświadczalnie. Dane uczące określają liczebność warstwy wejściowej – neuronów powinno być tyle, ile sygnałów wejściowych. Podobnie neuronów wyjściowych powinno być tyle, ile sygnałów wyjściowych. Ponieważ zbiór sygnałów wejściowych jest, co prawda, dyskretny, jednak o dużej liczbie możliwych stanów, określających cztery wielkości: miesiąc, dzień (numer dnia w miesiącu), dzień tygodnia i godzinę, przyjęto cztery neurony wejściowe. Na wyjściu otrzymujemy jedną wartość – cenę, która w ogólnym przypadku jest wartością rzeczywistą, wobec czego zastosowano 1 neuron na wyjściu. Liczebność warstwy ukrytej dobrano doświadczalnie. Z pewnością największy wpływ ma na jej rozmiar liczba danych trenujących. W rozpatrywanym przykładzie są to dane o cenach dla całego roku, czyli jest ich 365 razy 24, co daje 8760 danych uczących. W prezentowanych przykładach do uczenia wykorzystano dane pochodzące z 2011 roku. Przeprowadzone symulacje pokazały, że sieć jest w stanie nauczyć się poprawnie klasyfikować dane uczące (średni błąd to ok. 5-7%) przy 100 neuronach ukrytych. W trakcie nauki przedstawiano sieci neuronowej każdy wzorzec 100000 razy. Duża liczba epok uczenia wynika z przyjęcia bardzo małych współczynników uczenia. Przy większych wartościach tych współczynników proces nauki był bardzo chaotyczny i niestabilny z gwałtownymi spadkami, a następnie wzrostami błędów. Zmniejszenie wartości współczynników uczenia sieci uspokoiło proces nauki, choć również go znacząco spowolniło.

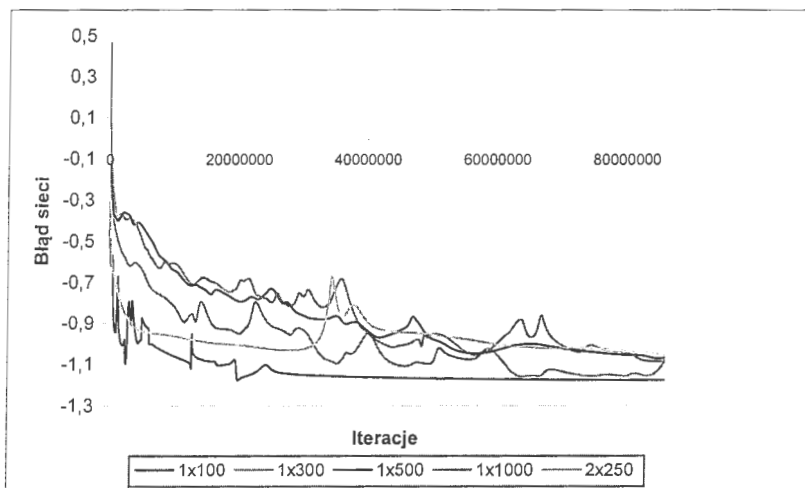
Aby zminimalizować możliwe działanie efektu tzw. przeuczenia sieci (sieć neuronowa może się bardzo dobrze nauczyć wzorców trenujących, lecz będzie niezbyt dobrze działać na danych testujących, pogarszając jeszcze ten wynik w trakcie dalszej nauki, mimo poprawy błędów dla danych trenujących) zastosowano uczenie ze sprawdzeniem co kilka epok uczących błędów występującego dla danych testujących. Zestaw wag z najniższym uzyskanym błędem dla danych testujących zostaje wybrany jako wynik nauki sieci.

4. Uczenie i wyniki działania sieci

4.1 Uczenie sieci

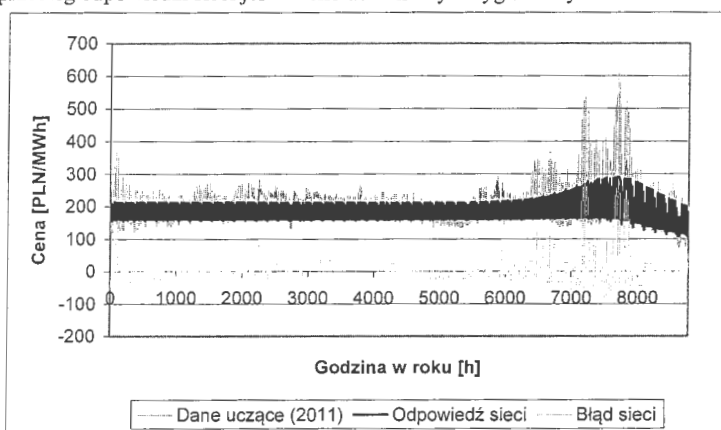
Samo uczenie sieci wygląda dość interesująco, bo w trakcie uczenia następuje znaczący spadek błędów, nawet dla niedużych sieci (rzędu 25 neuronów ukrytych i mniej). Jednak uczenie polega wtedy na znalezieniu w zasadzie jednej wartości, która jest wyliczana przez sieć, niezależnie od wejść. Oczywiście użyteczność takiego rozwiązania jest niewielka i świadczy o tym, że została przekroczona pojemność możliwości nauczania sieci. Na Rys 2. przedstawiono wykresy spadku błędów względnego na wyjściu sieci w zależności od jej wielkości dla dostatecznie dużych wartości, które nie powodują jej degradacji (niemożności

nauczenia wzorców). W legendzie Rys. 2. Podpisy typu „1x100” oznaczają sieć z jedną warstwą ukrytą o rozmiarze 100 neuronów, a „2x250” - sieć z dwiema warstwami ukrytymi o rozmiarze 250 neuronów każda. Na wykresie przedstawiono zlogarytmowany błąd uczenia, aby rysunek był bardziej czytelny.



Rys. 2. Przebieg zlogarytmowanego błędu uczenia sieci w zależności od liczby neuronów w warstwach ukrytych.

Jako efekt uczenia sieci można też przedstawić dane uczące, odpowiedź sieci oraz uzyskany błąd, co uwidoczniono jest na Rys. 3 dla sieci o wielkości 100 neuronów w jednej warstwie ukrytej. Uzyskano tu średni błąd względny nauki sieci o wartości 7% (sieć uczono zestawem danych nr 4). Analizując dane zamieszczone na rysunku można zauważyć, że sieć częściowo uwzględniła fluktuacje cen w końcu roku oraz tygodniową cykliczność cen, choć przebieg odpowiedzi sieci jest mocno uśredniony i wygładzony.



Rys. 3. Wykresy danych uczących (2011), odpowiedzi sieci i uzyskanego błędu dla sieci o 1 warstwie ukrytej o rozmiarze 100 neuronów .

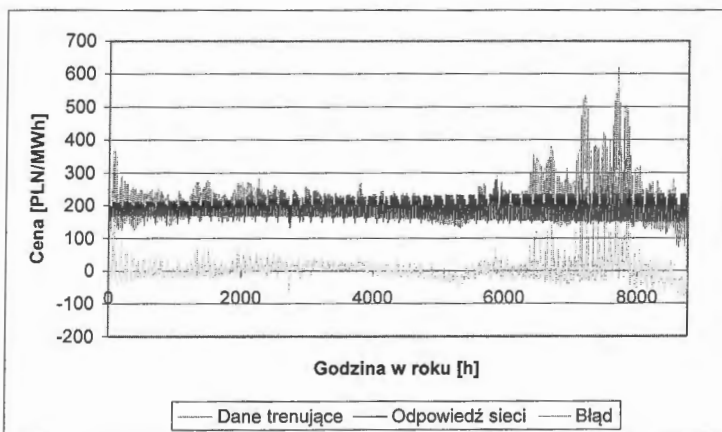
4.2 Wyniki działania sieci

W pierwszej fazie badań sprawdzono wielkość sieci, która pozwala uzyskać zadowalające parametry nauki i odpowiedzi na dane testujące. Wykonano szereg symulacji dla sieci od 20 do 1000 neuronów w 1 i 2 warstwach ukrytych. Na podstawie otrzymanych wyników i czasów uczenia sieci do dalszych doświadczeń wytypowana została sieć o 100 neuronach w 1 warstwie ukrytej. Mniejsze sieci uczyły się wprawdzie zdecydowanie szybciej, jednak ich nauka polegała na znalezieniu jednej wartości sygnału wyjściowego, która minimalizowała błąd właściwie bez zmian w zależności od podawanych sygnałów wejściowych. Sieci większe wymagały uczenia, które trwało nawet do tygodnia dla wersji o 1000 neuronów w 1 warstwie ukrytej. Sieć o 100 neuronach w 1 warstwie ukrytej uczy się w ciągu ok. 5 godzin z średnim błędem względnym ok. 7% (dla najlepszego z przedstawionych w punkcie 3.2 pięciu wariantów danych uczących) i średnim błędem względnym klasyfikacji danych testowych ok. 10%. Błąd nauki sieci może być mniejszy i spaść do 5%, jednak błąd klasyfikacji danych testowych jest wtedy większy – zaczyna być widoczny efekt przetrenowania sieci.

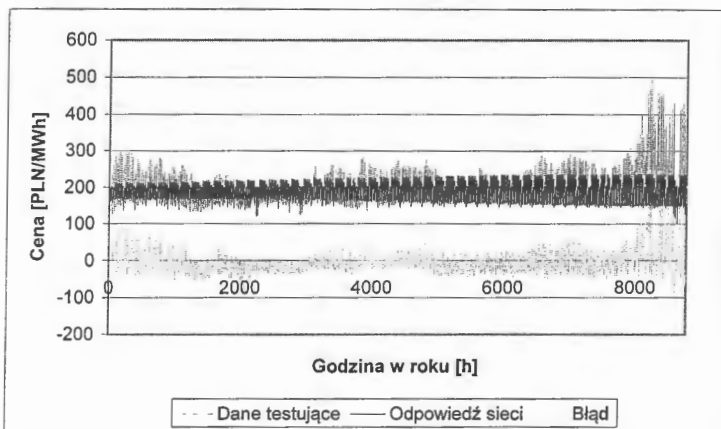
4.3. Wyniki nauki i działania sieci w zależności od użytego zestawu danych uczących.

Poniższe dane i rysunki przedstawiają wyniki nauki i rezultaty działania sieci w zależności od przyjętego zestawu danych uczących.

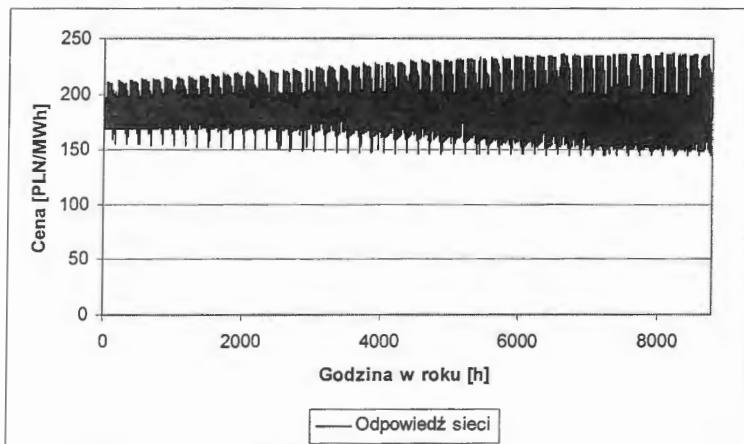
4.3.1. Sieć o wejściach: numer miesiąca, dzień miesiąca, dzień tygodnia, godzina doby i wskaźnik święta.



a) Dane trenujące (2011), odpowiedź sieci i błąd.



b) Dane testujące (2010), odpowiedź sieci i błąd.

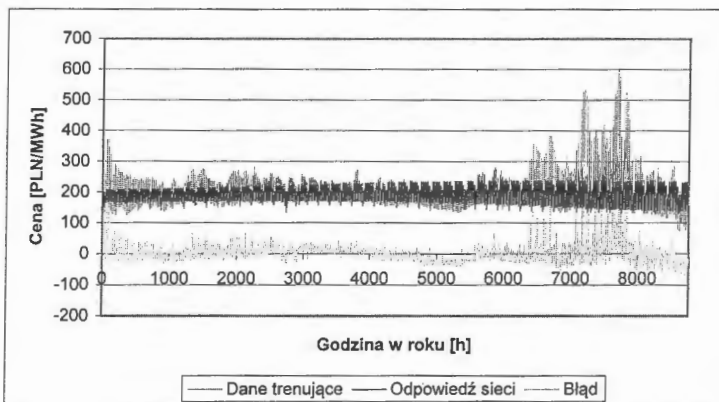


c) Odpowiedź sieci dla roku 2017.

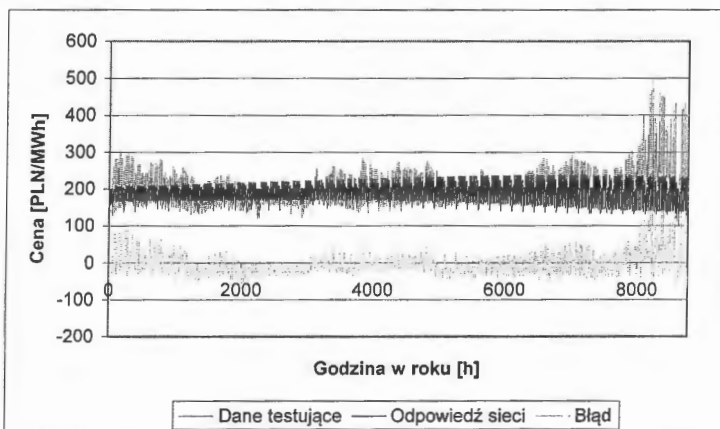
Rys. 4. Nauka i odpowiedź sieci o wejściach: numer miesiąca, dzień miesiąca, dzień tygodnia, godzina doby i wskaźnik święta.

Widoczne jest, że sieć, której wyniki działania przedstawia Rys. 4, nauczyła się danych dość schematycznie, reaguje jedynie na cykle dobowe i tygodniowe bez próby nauczenia się wyjątków, bądź dłuższych niż tydzień tendencji zmian danych uczących. Widoczna jest jedynie tendencja do wzrostu cen w ciągu roku, co z pewnością jest reakcją sieci na duże wartości cen energii w okresie przedświątecznym. Jednakże taka reakcja sieci bardzo słabo odzwierciedla tę tendencję. Średni błąd dla danych trenujących wynosi w tym przypadku 9,3%, a dla danych testujących 10,3%. Mimo niezbyt dużych błędów średnich takie rozwiązanie nie jest jednak wystarczająco dokładne, szczególnie dla okresu przedświątecznego..

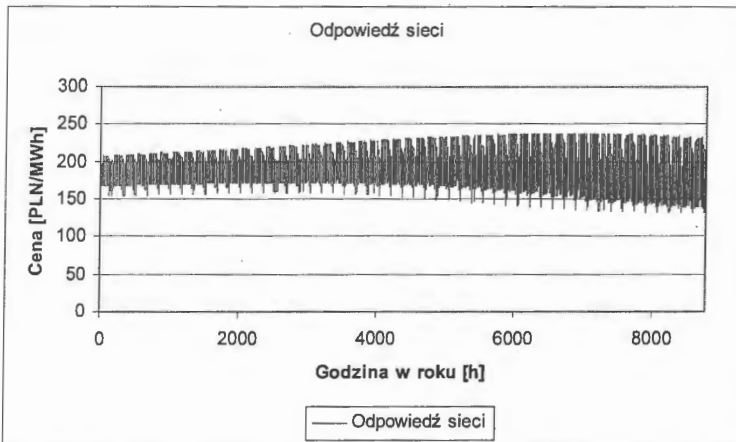
4.3.2. Sieć o wejściach: numer dnia w roku, dzień tygodnia, godzina doby i wskaźnik święta.



a) Dane trenujące (2011), odpowiedź sieci i błąd.



b) Dane testujące (2010), odpowiedź sieci i błąd.

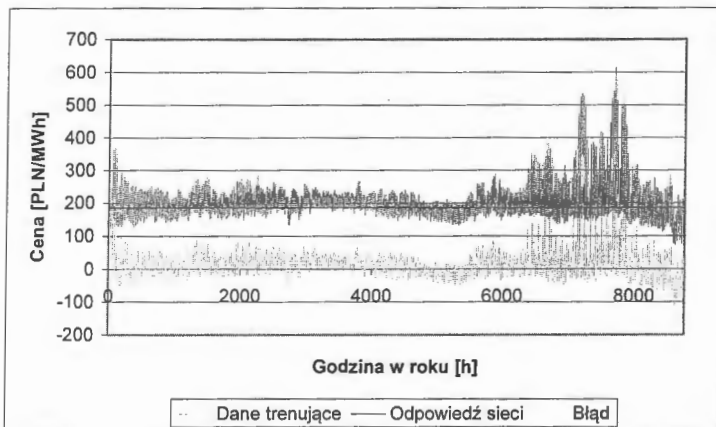


c) Odpowiedź sieci dla roku 2017.

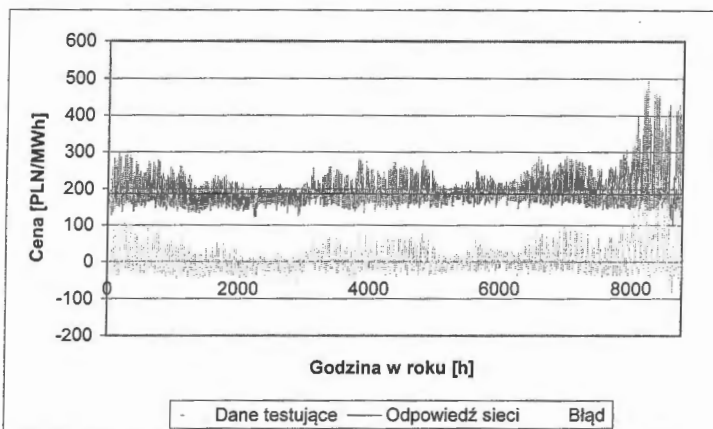
Rys. 5. Nauka i odpowiedź sieci o wejściach: numer dnia w roku, dzień tygodnia, godzina doby i wskaźnik święta.

Otrzymane w tym przypadku średnie błędy dla danych uczących to 9,2%, a dla danych testujących 10,4%. Ten zestaw danych zapewnia wyniki bardzo podobne do opisanych w poprzednim podpunkcie i również nie jest wystarczająco dokładny do zastosowania w systemie doradzania ceny jednostki energii.

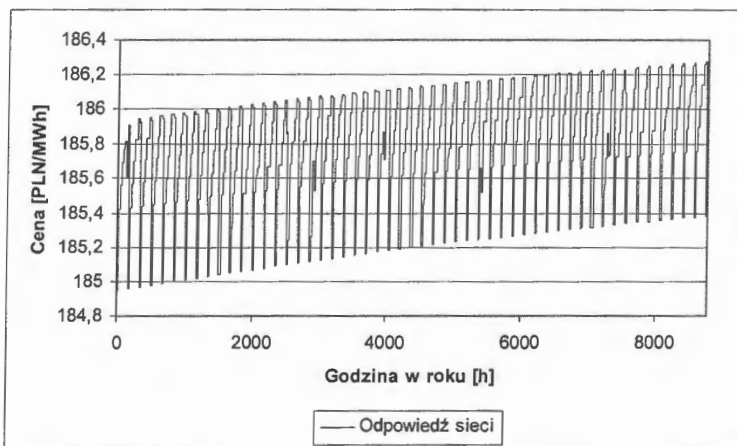
4.3.3. Sieć o wejściach: numer godziny w roku, dzień tygodnia i wskaźnik święta.



a) Dane trenujące (2011), odpowiedź sieci i błąd.



b) Dane testujące (2010), odpowiedź sieci i błąd.

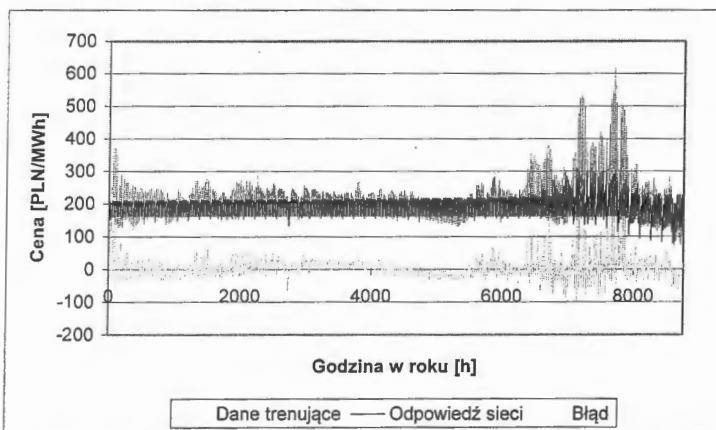


c) Odpowiedź sieci dla roku 2017.

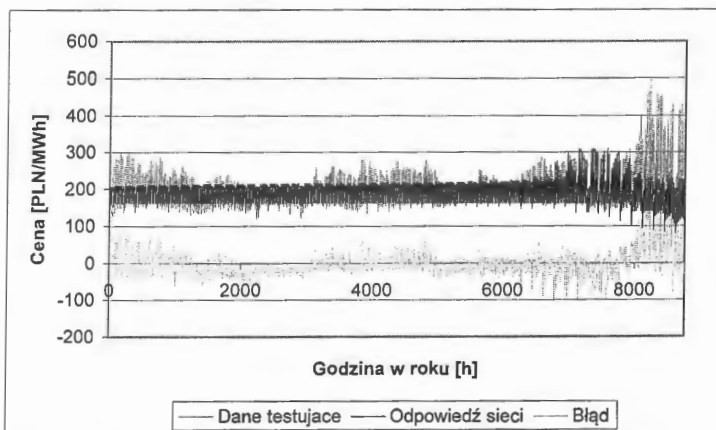
Rys. 6. Nauka i odpowiedź sieci o wejściach: numer godziny w roku, dzień tygodnia i wskaźnik święta.

Mimo, że uzyskane wartości błędu średniego nie są duże (13,7% dane trenujące i 12,3 dane testujące), to wyraźnie widać, że sieć o takich wejściach działa najgorzej z przebadanych i zdecydowanie nie nadaje się do użycia w celu przewidywania cen energii. Jest to związane ze zbyt małą różnorodnością danych na wejściu. Sieć nie potrafi sama powiązać danych przy tak ubogim zestawie danych wejściowych. Sieć reaguje bardzo zblizonymi odpowiedziami na dane wejściowe – przystosowała się do wartości średniej uczonego sygnału (zmienność wyjścia jest zbliżona do tej, widocznej na Rys. 6c, jednak nie widać tego na wykresie w przyjętej skali). Zaskakujący nieco lepszy wynik dla danych testujących niż trenujących jest zapewne przypadkowy.

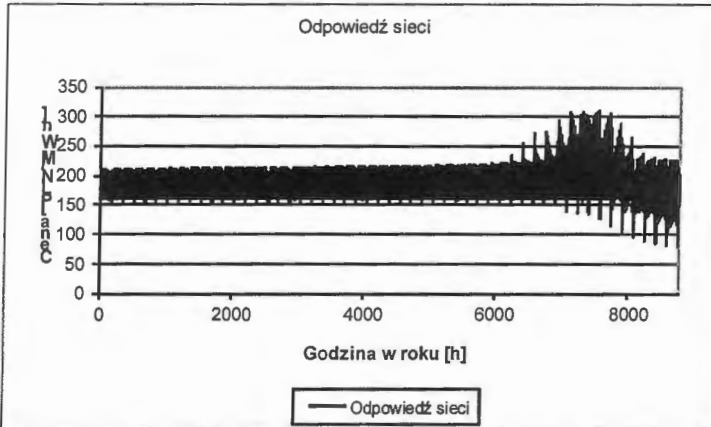
4.3.4. Sieć o wejściach: numer godziny w roku, dzień tygodnia, godzina doby i wskaźnik święta.



a) Dane trenujące (2011), odpowiedź sieci i błąd.



b) Dane testujące (2010), odpowiedź sieci i błąd.

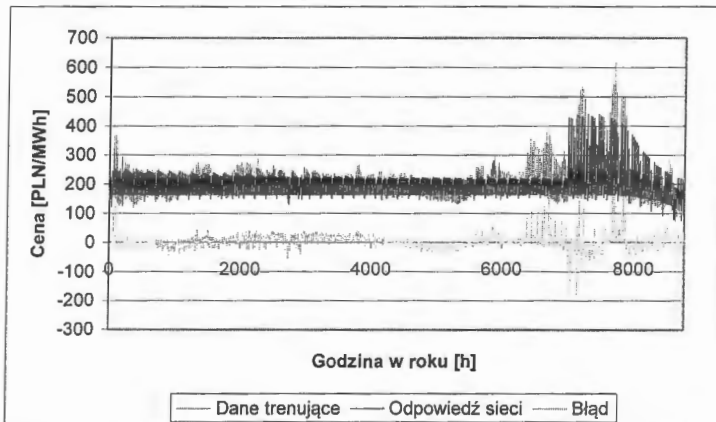


c) Odpowiedź sieci dla roku 2017.

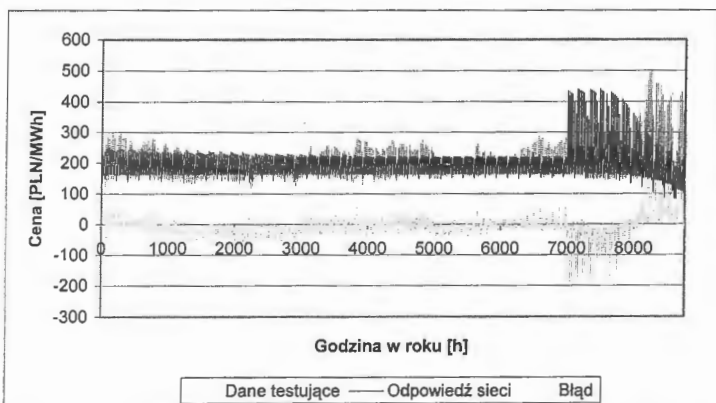
Rys. 7. Nauka i odpowiedź sieci o wejściach: numer godziny w roku, dzień tygodnia, godzina doby i wskaźnik święta.

Przedstawiana sieć zdecydowanie dobrze uczy się przebiegu rocznego i wykrywa wzrost, a następnie spadek cen pod koniec roku. Otrzymane błędy wynoszą odpowiednio: 8,4% dla danych trenujących i 10,5% dla danych testujących. Mimo, że sieć dobrze wykrywa wahania cen w okresie przedświątecznym, to jednak widać, że wyuczony przebieg tego zjawiska dla roku 2011 jest inny niż w roku 2010. Sieć odpowiada podobnie do wyuczonego roku 2011. Jest to oczywiście zgodne z oczekiwaniami, gdyż sieć uczono tylko na podstawie przebiegu z roku 2011. Jak widać kształty wahań w okresie przed- i po- świątecznym są różne w różnych latach. Być może ma na to jakiś wpływ pogoda. Jednak do pełniejszego uwzględnienia tego zjawiska konieczne byłoby uczenie sieci danymi z wielu lat. Jest to oczywiście możliwe, jednak czas potrzebny na nauczenie sieci wzrośnie wtedy wielokrotnie.

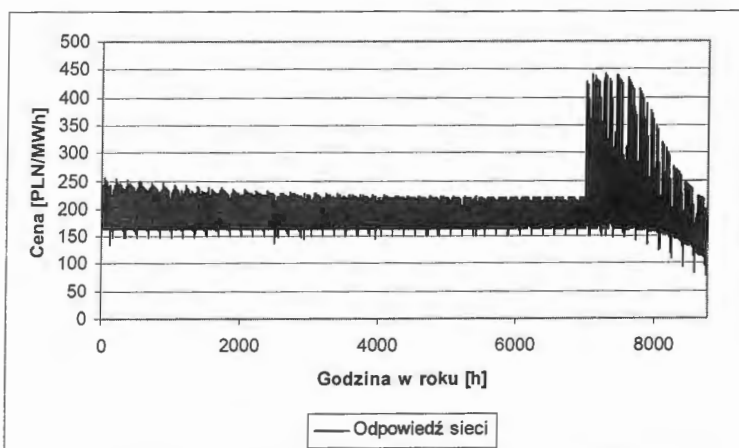
4.3.5. Sieć o wejściach: numer godziny w roku, dzień tygodnia, godzina doby i wskaźnik święta oraz identyfikator okresu przedświątecznego



a) Dane trenujące (2011), odpowiedź sieci i błąd.



b) Dane testujące (2010), odpowiedź sieci i błąd.

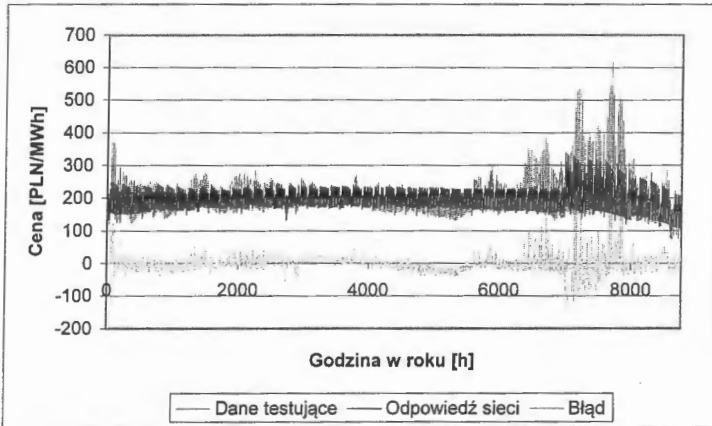


c) Odpowiedź sieci dla roku 2017.

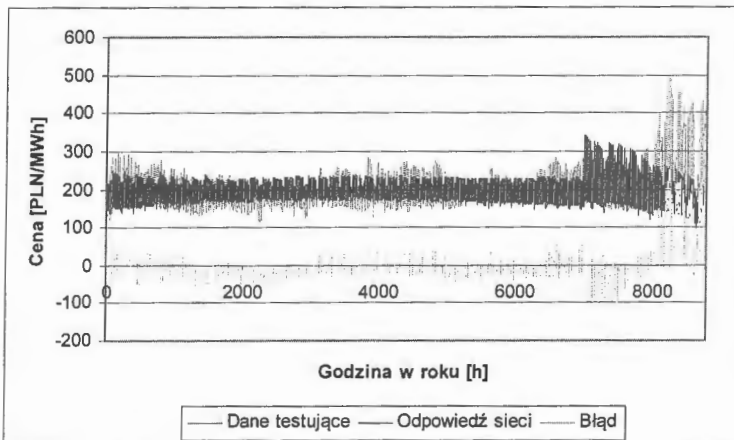
Rys. 8. Nauka i odpowiedź sieci o wejściach: numer godziny w roku, dzień tygodnia, godzina doby i wskaźnik święta oraz identyfikator okresu przedświątecznego.

Dołączenie do zestawu uczącego sygnału informującego sieć o okresie przedświątecznym ułatwiło sieci wykrycie tej sytuacji, jednakże odpowiedź sieci, szczególnie dla prognozy w roku 2017, wygląda na znacząco przesterowaną. Błąd uzyskany podczas uczenia sieci wynosi 7,2%, a podczas testowania 9,5%. Jest to wynik najlepszy z uzyskanych we wszystkich badaniach.

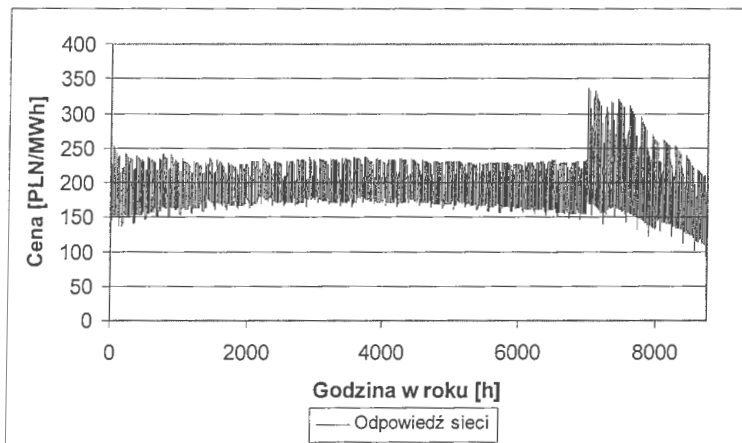
4.3.6. Sieć o wejściach: numer dnia w roku, numer miesiąca, numer dnia w miesiącu, numer godziny w roku, dzień tygodnia, godzina doby i wskaźnik święta oraz identyfikator okresu przedświątecznego



a) Dane trenujące (2011), odpowiedź sieci i błąd.



b) Dane testujące (2010), odpowiedź sieci i błąd.



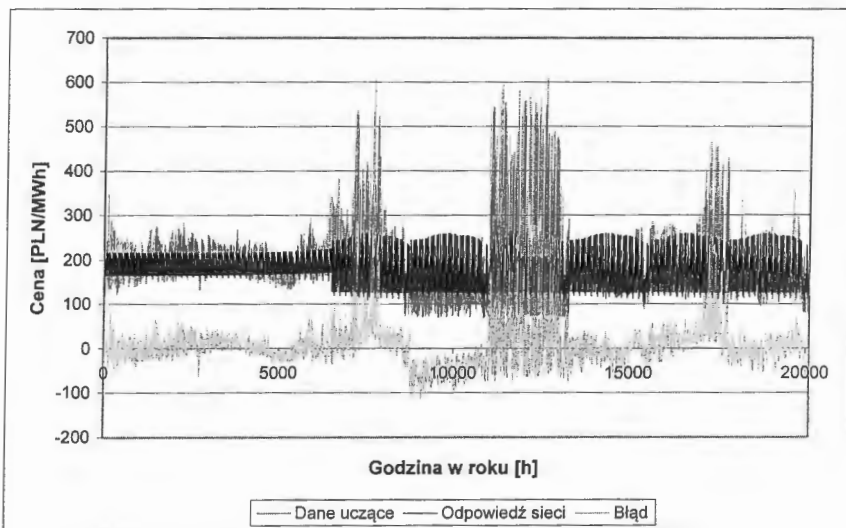
c) Odpowiedź sieci dla roku 2017.

Rys. 9. Nauka i odpowiedź sieci o wejściach: numer dnia w roku, numer miesiąca, numer dnia w miesiącu, numer godziny w roku, dzień tygodnia, godzina doby i wskaźnik święta oraz identyfikator okresu przedświątecznego.

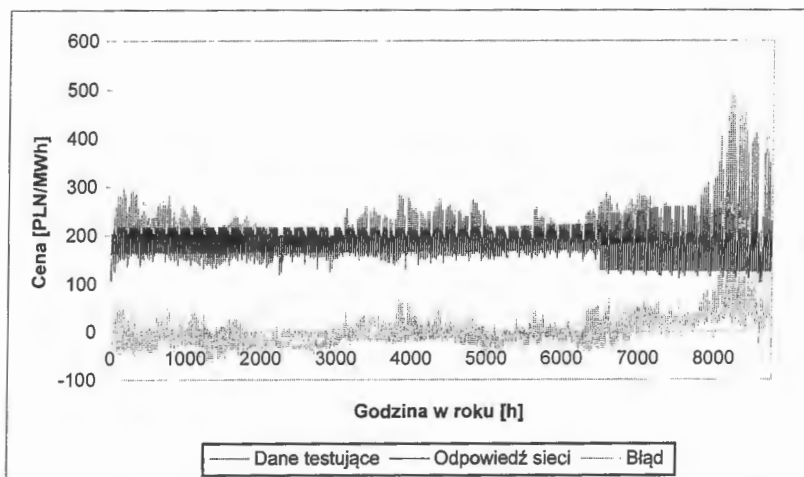
W tym przypadku błąd uzyskany podczas uczenia sieci wynosi 7,3%, a podczas testowania 10,6%, czyli nieco gorzej niż w przypadku opisywanym w punkcie poprzednim. Ogólnie rzecz biorąc wyniki otrzymane w punktach 4.3.5. i 4.3.6 są bardzo podobne i mogą być wykorzystane w module prognozującym ceny energii.

4.3.7. Sieć o zwielokrotnionych danych o końcu roku.

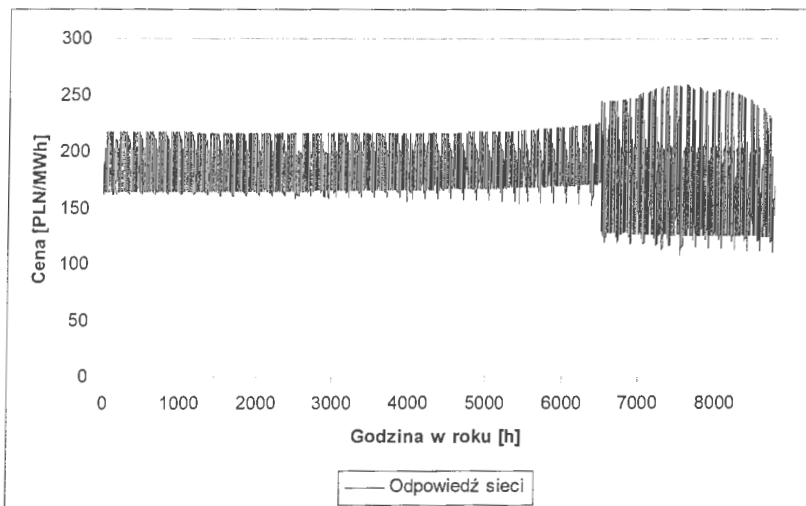
W opisywanych wyżej przypadkach dość problemy sprawiał koniec roku, w którym ceny energii nadmiernie rosną, prawdopodobnie w związku z różnego rodzaju przygotowaniem świątecznymi oraz dodatkowymi działaniami firm, związanymi z końcem roku. Niestety kształt krzywej cen i jej wartości dość znacznie różnią się w poszczególnych latach. Podjęto wobec tego próbę uwzględnienia tego faktu w danych uczących, przez dodanie do danych uczących z roku 2011 dodatkowych danych z końców lat 2007-2009 oraz 2012 (rok 2010 jest wykorzystywany jako dane testujące). W sumie daje to 20110 danych trenujących. Wyniki działania takiego modelu z zastosowaniem następujących wejść: numer godziny w roku, dzień tygodnia, godzina doby i wskaźnik święta oraz identyfikator okresu przedświątecznego, przedstawiają zamieszczone dalej rysunki.



a) Dane trenujące (2011 oraz 2007-2012), odpowiedź sieci i błąd.



b) Dane testujące (2010), odpowiedź sieci i błąd.



Rys. 10. Nauka i odpowiedź sieci o wejściach: numer godziny w roku, dzień tygodnia, godzina doby, wskaźnik święta oraz identyfikator okresu przedświątecznego.

W prezentowanych przypadkach dokładność nauki danych wynosi 16%, a dokładność klasyfikacji danych testujących wynosi 10,8 %. Ta nieco nietypowa sytuacja związana jest z wykorzystaniem dużo większego i bardziej skomplikowanego zbioru danych trenujących niż danych testujących.

5. Wnioski

Sieć neuronowa typu backpropagation jest narzędziem, które z powodzeniem można wykorzystać do prognozowania cen na rynku energii. Jednakże, aby polepszyć dokładność działania sieci, z pewnością konieczne jest uwzględnienie danych z kilku lat. Powoduje to pewne problemy. Co prawda posiadane dane, zawierają informacje z 12 lat, jednakże dane te w większości zawierają dużo „zer”, czyli pozycji, w których nie odbyły się transakcje. Wartości zerowe nie mogą być wykorzystywane do nauki sieci, gdyż bardzo zaburzają ten proces, a z kolei interpolowanie wyników powoduje wygładzanie danych, przez co pozbywają się one istotnych informacji. Dane z małą liczbą „zer” występują dla 4 zestawów rocznych, czyli nie jest już ich tak dużo.

Uwzględnienie roku w danych trenujących znacząco zwiększy ich liczbę, co spowoduje konieczność zwiększenia liczby neuronów, koniecznych do „zapamiętania” większej liczby danych, a to z kolei znacząco wydłuży proces uczenia. Biorąc pod uwagę doświadczenia z obecną postacią sieci, należy sądzić, że czas nauki może wzrosnąć nawet do kilku tygodni nieprzerwanej pracy komputera. Oczywiście w przypadkach praktycznego wykorzystania takiego modułu kilkutygodniowe uczenie sieci może być akceptowalne.

Literatura

- [1] A. N. Kolmogorov: On the representation of continuous functions of many variables by superposition of continuous functions of one variable and addition. *Dokl. Akad. Nauk USSR*, 114: 953–956, 1957.
- [2] J. Hertz, A. Krogh, R. G. Palmer: *Wstęp do teorii obliczeń neuronowych*. WNT, Warszawa, 1993.
- [3] J. Korbicz, A. Obuchowicz, D. Uciński: *Sztuczne sieci neuronowe. Podstawy i zastosowania*. Akademicka Oficyna Wydawnicza PLJ, Warszawa, 1994.
- [4] R. Hecht-Nielsen: *Neurocomputing*. Addison-Wesley, 1990.

the 1990s, the number of people with diabetes has increased in all industrialized countries. In the Netherlands, the prevalence of diabetes has increased from 1.5% in 1975 to 5.5% in 1995 (1). The prevalence of diabetes is expected to increase further in the next decades (2).

Diabetes is a chronic disease with a high prevalence of complications. The most common complications are retinopathy, nephropathy, neuropathy, and cardiovascular disease. The prevalence of these complications is high, and the risk of complications is increased in people with diabetes (3). The prevalence of complications is also increased in people with diabetes who are treated with insulin (4).

The prevalence of complications is also increased in people with diabetes who are treated with oral hypoglycaemic agents (5). The prevalence of complications is also increased in people with diabetes who are treated with insulin and oral hypoglycaemic agents (6). The prevalence of complications is also increased in people with diabetes who are treated with insulin and oral hypoglycaemic agents (7).

The prevalence of complications is also increased in people with diabetes who are treated with insulin and oral hypoglycaemic agents (8). The prevalence of complications is also increased in people with diabetes who are treated with insulin and oral hypoglycaemic agents (9). The prevalence of complications is also increased in people with diabetes who are treated with insulin and oral hypoglycaemic agents (10).

The prevalence of complications is also increased in people with diabetes who are treated with insulin and oral hypoglycaemic agents (11). The prevalence of complications is also increased in people with diabetes who are treated with insulin and oral hypoglycaemic agents (12). The prevalence of complications is also increased in people with diabetes who are treated with insulin and oral hypoglycaemic agents (13).

The prevalence of complications is also increased in people with diabetes who are treated with insulin and oral hypoglycaemic agents (14). The prevalence of complications is also increased in people with diabetes who are treated with insulin and oral hypoglycaemic agents (15). The prevalence of complications is also increased in people with diabetes who are treated with insulin and oral hypoglycaemic agents (16).

The prevalence of complications is also increased in people with diabetes who are treated with insulin and oral hypoglycaemic agents (17). The prevalence of complications is also increased in people with diabetes who are treated with insulin and oral hypoglycaemic agents (18). The prevalence of complications is also increased in people with diabetes who are treated with insulin and oral hypoglycaemic agents (19).

The prevalence of complications is also increased in people with diabetes who are treated with insulin and oral hypoglycaemic agents (20). The prevalence of complications is also increased in people with diabetes who are treated with insulin and oral hypoglycaemic agents (21). The prevalence of complications is also increased in people with diabetes who are treated with insulin and oral hypoglycaemic agents (22).

the 1990s, the number of people in the world who are living in poverty has increased from 1.2 billion to 1.6 billion (World Bank 2000).

There are a number of reasons for this increase in poverty. One of the main reasons is the rapid growth of the world population. The world population is expected to reach 8 billion by the year 2025 (United Nations 2000). This increase in population has led to a corresponding increase in the demand for food, water, and other resources. As a result, the price of these resources has increased, and this has led to an increase in the cost of living for many people in the world.

Another reason for the increase in poverty is the rapid growth of the world economy. The world economy has grown rapidly in the 1990s, and this has led to an increase in the number of people who are living in poverty. This is because the growth of the world economy has led to an increase in the number of people who are working in the informal sector of the economy. These people are often paid very low wages and do not have any benefits or job security.

A third reason for the increase in poverty is the rapid growth of the world's debt. The world's debt has increased rapidly in the 1990s, and this has led to an increase in the number of people who are living in poverty. This is because the increase in the world's debt has led to an increase in the interest payments that countries have to make. This has led to a decrease in the amount of money that countries have available for social services, and this has led to an increase in the number of people who are living in poverty.

There are a number of ways in which the world can reduce the number of people who are living in poverty. One of the most important ways is to reduce the world's population. This can be done by increasing the age at which people have children, and by increasing the number of children that people have. This will lead to a decrease in the world's population, and this will lead to a decrease in the demand for food, water, and other resources. As a result, the price of these resources will decrease, and this will lead to a decrease in the cost of living for many people in the world.

Another way in which the world can reduce the number of people who are living in poverty is to reduce the world's debt. This can be done by increasing the amount of money that countries have available for social services, and by increasing the amount of money that countries have available for investment. This will lead to an increase in the number of people who are working in the formal sector of the economy, and this will lead to an increase in the number of people who are living in poverty.

There are a number of other ways in which the world can reduce the number of people who are living in poverty. These include increasing the minimum wage, increasing the number of people who are working in the informal sector of the economy, and increasing the number of people who are receiving social services. All of these measures will lead to a decrease in the number of people who are living in poverty.