

Inspirowane kwantowo sieci neuronowe typu Hopfielda

Streszczenie. W pracy przedstawiono koncepcję sieci neuronowej o zespolonych parametrach (Q-inspired). Realizacja takiej sieci wykorzystuje hermitowską macierz połączeń pomiędzy neuronami. Zaproponowano również model uczenia maszynowego zrealizowany na bazie zespolonego aproksymatora. Wykazano przydatność takiego aproksymatora w analizie sygnałów w szczególności do realizacji dyskretnej transformacji Fouriera (DFT) oraz odwrotnej dyskretnej transformacji Fouriera (IDFT).

Abstract. The paper presents the concept of a neural network with complex-valued parameters (Q-inspired). Implementation of such a network uses the Hermitian matrix of connections between neurons. A machine learning model based on a complex approximator has also been proposed. The usefulness of such an approximator in signal analysis has been demonstrated especially for the implementation of discrete Fourier transform (DFT) and inverse discrete Fourier transform (IDFT). (**Q-inspired Hopfield Type Neural Networks**).

Słowa kluczowe: inteligencja obliczeniowa, uczenie maszynowe, kwantowo inspirowane sieci neuronowe, DFT.

Keywords: artificial intelligence, machine learning, Q-inspired neural networks, DFT.

Wstęp

Sztuczne sieci neuronowe stanowią uniwersalne modele algorytmiczne i fizyczne wykorzystywane w dziedzinie inteligencji obliczeniowej. Dotychczas nie ustalono optymalnej architektury i technologii realizacji takich sieci. Można natomiast zaobserwować, że badania nad takimi modelami prowadzone są w trzech zasadniczych kierunkach. Są to tzw. klasyczne modele obliczeniowe (nie kwantowe) o rzeczywistych parametrach (RVNN-Real Valued Neural Networks), modele klasyczne o zespolonych parametrach (CVNN – Complex Valued Neural Networks) oraz kwantowe sieci neuronowe (QNN-Quantum Neural Networks), stanowiące pewną alternatywę, dla uniwersalnych komputerów kwantowych (QC - Quantum Computers), w rozwiązywaniu problemów należących do dziedziny sztucznej inteligencji. Należy zauważyć, że tzw. kwantowo inspirowane (Q-inspired) sieci są pewną wersją modeli o zespolonych parametrach. Są to jednak struktury, klasyczne nie kwantowe[1, 2]. Wśród struktur o rzeczywistych parametrach istotną rolę odgrywają układy neuronowe Hopfielda. W poprzednich pracach [3, 4, 5] pokazano jak rozszerzając macierz połączeń układów Hopfielda o składowe antysymetryczne, można uzyskać uniwersalne modele, do rozwiązywania zagadnień z dziedziny nauczania maszynowego. Celem niniejszej pracy jest opis sieci neuronowej o zespolonych parametrach (Q-inspired) uzyskanej poprzez zastąpienie macierzy symetrycznych w/w modelach, macierzami hermitowskimi. Pozwala to na zwiększoną efektywność obliczeniową modelu nauczania maszynowego.

Rozszerzone sieci typu Hopfielda

Rozszerzone sieci typu Hopfielda są pewnym uogólnieniem klasycznych hopfieldowskich struktur. Można je opisać wektorowym równaniem różniczkowym o postaci:

$$(1) \quad \dot{\mathbf{x}} = (\eta \mathbf{W} - w_0 \mathbf{1} + \varepsilon \mathbf{W}_s) \boldsymbol{\theta}(\mathbf{x}) + \mathbf{I}_d$$

gdzie: \mathbf{W} - macierz ortogonalna, skośnie symetryczna, \mathbf{W}_s - macierz rzeczywista symetryczna, $\mathbf{1}$ - macierz jednostkowa, $\boldsymbol{\theta}(\mathbf{x})$ – wektor funkcji aktywacji: $\mu_1 \leq \frac{\theta(x)}{x} \leq \mu_2$; $\mu_1, \mu_2 \in (0, \infty)$, \mathbf{I}_d – wektor wejściowy, ε, w_0, η - parametry rzeczywiste.

Przedstawione powyżej równanie definiuje pewną ogólną klasę sieci typu Hopfielda, która w zależności od doboru parametrów ε, w_0, η , występujących w równaniu, można podzielić na pewne podtypy:

- Bezstratna sieć hamiltonowska: $\varepsilon, w_0 = 0, \eta = 1$.
- Sieć Hopfielda: $w_0, \eta = 0$.
- Kwantowo inspirowana sieć typu Hopfielda. Sieć taka należy do układów o zespolonych parametrach i stanowi pewną algebraiczną transformację o postaci:

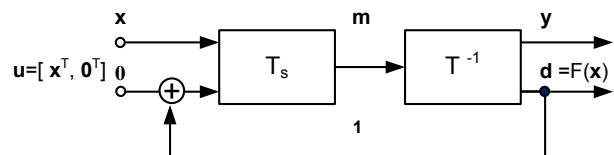
$$(2) \quad (\eta \mathbf{W} - w_0 \mathbf{1} + \varepsilon \mathbf{W}_H) \boldsymbol{\theta}(\mathbf{x}) + \mathbf{I}_d = \mathbf{0}$$

gdzie: \mathbf{W}_H - macierz hermitowska ($\mathbf{W}_H = \mathbf{W}_H^\dagger$, operator † oznacza sprzężenie hermitowskie).

W przypadku gdy \mathbf{W}_H przyjmuje wartości rzeczywiste, równanie (2) jest równaniem równowagi rozszerzonej sieci Hopfielda.

Struktura modelu uczenia maszynowego

Model uczenia maszynowego uzyskano przez połączenie w pierścieni sieci opisane równaniem (2). Strukturę rozważanego modelu uczenia maszynowego pokazano na rysunku 1. Jego podstawową funkcją jest implementacja pewnego odwzorowania $\mathbf{d} = F(\mathbf{x})$, zadanego w punktach treningowych $\{\mathbf{x}_i, \mathbf{d}_i\}_{i=1}^N$.



Rys. 1. Struktura modelu uczenia maszynowego

Dla wspomnianego zbioru treningowego $S = \{\mathbf{x}_i, \mathbf{d}_i\}_{i=1}^N$, gdzie $\mathbf{x}_i \in \mathbb{C}^n, \mathbf{d}_i \in \mathbb{C}^m$ generuje się zbiór wektorów systemowych \mathbf{u}_i o postaci:

$$(3) \quad \mathbf{u}_i = \begin{bmatrix} \mathbf{x}_i \\ \mathbf{d}_i \end{bmatrix}, i = 1, \dots, N$$

gdzie: $\dim \mathbf{u}_i = n + m, n + m = 2^k, k = 2, 3, \dots$

Transformacja ortogonalna $T(\cdot)$, opierająca się na rodzinie macierzy Hurwitza-Radona, pozwala na wyznaczenie widma Haar'a-Fouriera wektorów systemowych $\mathbf{u}_i, i = 1, \dots, N$:

$$(4) \quad \mathbf{m}_i = \frac{1}{2} (\mathbf{W}_{2^k} + \mathbf{1}) \mathbf{u}_i = T(\mathbf{u}_i)$$

gdzie: \mathbf{W}_{2^k} – macierz Hurwitza-Radona, $\mathbf{W}_{2^k}^2 = -\mathbf{1}$, \mathbf{m}_i – widma wektorów $\mathbf{u}_i, i = 1, \dots, N$, $\mathbf{M} = \{\mathbf{m}_1, \mathbf{m}_2, \dots, \mathbf{m}_N\}$ – macierz zespolona wektorów widma, $\mathbf{1}$ - macierz jednostkowa.

Blok T^{-1} systemu z rysunku 1 realizuje operację:

$$(5) \quad \mathbf{u}_i = (-W_{2^k} + \mathbf{1})\mathbf{m}_i = T^{-1}(\mathbf{m}_i)$$

Natomiast blok T_s transformacją biortogonalną:

$$(6) \quad \mathbf{m}_i = (2 \cdot \mathbf{1} - W_H - W_{2^k})^{-1} \mathbf{u}_i$$

Stąd

$$(7) \quad T_s(\cdot) = (2 \cdot \mathbf{1} - W_H - W_{2^k})^{-1}$$

gdzie: $W_H = \mathbf{M}(\mathbf{M}^T \mathbf{M})^{-1} \mathbf{M}^T$ – macierz hermitowska

Struktura z rysunku 1 pełni funkcję estymatora widma $\{\hat{\mathbf{m}}_i\}$:

$$(8) \quad \hat{\mathbf{m}}_i = T_s \left(\begin{bmatrix} \mathbf{x}_i \\ \mathbf{0} \end{bmatrix} \right), i = 1, \dots, N$$

gdzie: $\mathbf{0}$ – wektor zerowy.

W systemie, ze względu na działanie pętli sprzężenia zwrotnego, uzyskuje się konwergencję wektorów:

$$(9) \quad \hat{\mathbf{m}}_i \rightarrow \mathbf{m}_i, \hat{\mathbf{y}}_i \rightarrow \mathbf{x}_i$$

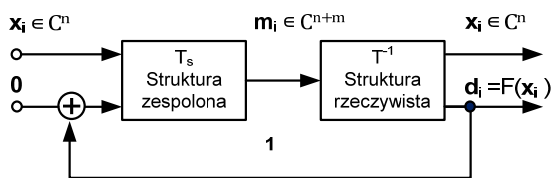
oraz sygnału wyjściowego:

$$(10) \quad \hat{\mathbf{d}}_i \rightarrow \mathbf{d}_i = F(\mathbf{x}_i), i = 1, \dots, N$$

Można także zauważyć, że każda superpozycja $\mathbf{x}_s = \sum_{k=1}^N \alpha_k \mathbf{x}_k$ wektorów przynależnych do zbioru treningowego implikuje superpozycję wektorów wyjściowych $\sum_{k=1}^N \alpha_k \mathbf{d}_k$, $\alpha_k \in \mathbb{C}$. Tę własność wykorzystuje się w projekcie procesora analogowego [4].

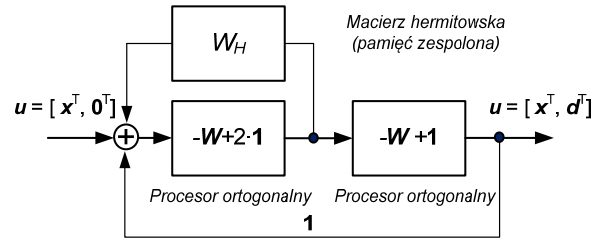
Własności inspirowanego kwantowo modelu uczenia maszynowego

Modele uczenia maszynowego oparte na hamiltonowskich sieciach neuronowych realizuje się wykorzystując macierze połączeń z rzeczywistymi współczynnikami. Okazuje się jednak, że można takie struktury łatwo rozszerzyć realizując macierze połączeń o współczynnikach zespolonych. Takie podejście zwiększa funkcjonalność badanych struktur i zwiększa ich efektywność obliczeniową. Tak więc dla zespolonego zbioru wektorów treningowych $\{\mathbf{x}_i, \mathbf{d}_i\}_{i=1}^N$, gdzie $\mathbf{x} \in \mathbb{C}^n$, $\mathbf{d}_i \in \mathbb{C}^m$, $n + m = 2^k, k = 3, 4, \dots$ stosując wcześniej wspomnianą transformację biortogonalną uzyskuje się przekształcenie zespolone: $\mathbb{C}^n \rightarrow \mathbb{C}^m$. Struktura aproksymatora realizującego takie przekształcenie pokazana jest na rysunku 2.



Rys.2. Struktura zespolonego aproksymatora z rozłożoną pamięcią

Przedstawiony aproksymator może zostać zrealizowany jako Q-inspirowana sieć neuronowa z macierzą połączeń o wartościach zespolonych. Sieć ta odwzorowuje algebraicznie przestrzeń zespoloną \mathbb{C}^n w przestrzeń zespoloną \mathbb{C}^m : $\mathbb{C}^n \rightarrow \mathbb{C}^m$. Schemat sieci ze skupioną pamięcią przedstawiono na rysunku 3.



Rys.3. Schemat blokowy zespolonego aproksymatora ze skupioną pamięcią

Doświadczenia numeryczne pokazują, że efektywność obliczeniowa zespolonego aproksymatora jest lepsza niż zrealizowanego w oparciu o macierze z rzeczywistymi współczynnikami. Zastosowanie zespolonego aproksymatora pozwala na przykład na realizację dyskretnego przekształcenia Fouriera (DFT). Zbiór wektorów treningowych $S = \{\mathbf{x}_i, \mathbf{d}_i\}_{i=1}^N$ w takim przypadku jest generowany za pomocą unitarnej transformacji DFT:

$$(11) \quad DFT\{\mathbf{x}_i\} = \mathbf{d}_i, i = 1, \dots, N$$

gdzie: $\dim \mathbf{x}_i = m$, $\dim \mathbf{d}_i = m$, $m = 2^k, k = 3, 4, \dots$, $\mathbf{d}_i \in \mathbb{C}^m$, m – punktowe widmo $\mathbf{x}_i (m > 8)$

Dysponując m liniowo niezależnymi wektorami $\mathbf{z}_1, \dots, \mathbf{z}_m$, $\dim \mathbf{z}_i = m$, $\mathbf{z}_i \in \mathbb{R}^m$ możemy uzyskać dowolny wektor \mathbf{x}_i jako superpozycję:

$$(12) \quad \mathbf{x}_i = \alpha_1 \mathbf{z}_1 + \dots + \alpha_m \mathbf{z}_m, \mathbf{x}_i \in S$$

$$(13) \quad \boldsymbol{\alpha} = [\alpha_1, \dots, \alpha_m]^T = \mathbf{Z}^{-1} \mathbf{x}_i$$

gdzie macierz $\mathbf{Z} = [\mathbf{z}_1, \dots, \mathbf{z}_m]$ nie jest macierzą jednostkową.

W takim wypadku struktura aproksymatora jest generowana poprzez przekształcenie:

$$(14) \quad \mathbf{u}_i = \begin{bmatrix} \mathbf{z}_i \\ DFT(\mathbf{z}_i) \end{bmatrix}; i = 1, \dots, m, \dim \mathbf{u}_i = 2m$$

oraz

$$(15) \quad \phi \begin{bmatrix} \mathbf{z}_i \\ \mathbf{0} \end{bmatrix} = \begin{bmatrix} \mathbf{z}_i \\ DFT(\mathbf{z}_i) \end{bmatrix}, i = 1, \dots, m$$

W celu zapewnienia stabilności numerycznej aproksymatora musi być spełniony warunek:

$$(16) \quad N < 0.5 (2m) = m$$

W związku z powyższym do realizacji $DFT\{\mathbf{x}_i\}, i = 1, \dots, L$ potrzebujemy co najmniej dwóch aproksymatorów. Każdy wektor $\mathbf{x}_i \in S$ możemy przedstawić jako sumę dwóch wektorów składowych:

$$(17) \quad \mathbf{x}_i = \mathbf{x}_{1i} + \mathbf{x}_{2i}, i = 1, \dots, L$$

gdzie: $\mathbf{x}_{1i} = \alpha_1 \mathbf{z}_1 + \dots + \alpha_l \mathbf{z}_l$, $\mathbf{x}_{2i} = \alpha_{l+1} \mathbf{z}_{l+1} + \dots + \alpha_m \mathbf{z}_m$, $l = m/2$

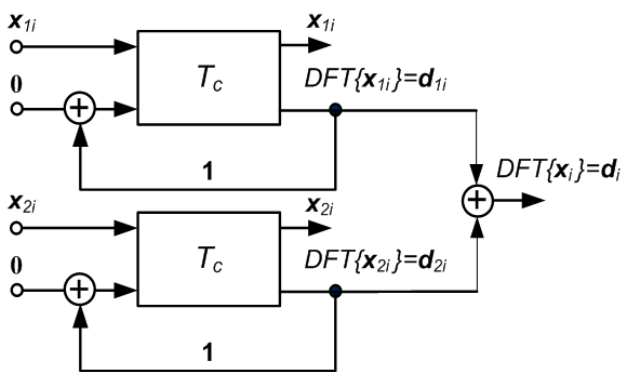
oraz

$$(18) \quad DFT\{\mathbf{x}_i\} = DFT\{\mathbf{x}_{1i}\} + DFT\{\mathbf{x}_{2i}\}.$$

Korzystając z powyższych założeń uzyskujemy strukturę dwóch aproksymatorów wyznaczających transformatę DFT:

$$(19) \quad \mathbf{u}_i = \begin{bmatrix} \mathbf{z}_i \\ DFT(\mathbf{z}_i) \end{bmatrix}; i = 1, \dots, l$$

$$(20) \quad \mathbf{u}_i = \begin{bmatrix} \mathbf{z}_i \\ DFT(\mathbf{z}_i) \end{bmatrix}; i = l + 1, \dots, m, l = m/2.$$



Rys.4. Schemat blokowy procesora DFT

Aproksymator pokazany na rysunku 4 oblicza widmo DFT wektora d_i dla dowolnego wektora $x_i \in S$ a ponadto oblicza widmo d dla dowolnego wektora $x < \infty$ ($\dim x = m$). W tym samym systemie możliwe jest modelowanie odwrotne, istnieje zatem możliwość realizacji odwrotnej transformacji Fouriera (IDFT). Warto także zauważyć, że opisany w pracy Q-inspirowany model uczenia maszynowego może być używany jako model procesora analogowego do dodawania i mnożenia liczb zespolonych.

Podsumowanie

W pracach [3, 4, 5] przedstawiono schemat uczenia maszynowego, wykorzystujący transformację ortogonalną i biotrogonalną, zrealizowany w oparciu o rozszerzony układ sieci Hopfielda. Podstawową funkcją systemu jest realizacja odwzorowania, zdefiniowanego w przestrzeni liczb rzeczywistych, $d = F(x)$, $d \in R^m, x \in R^n$. Odwzorowanie generowane jest w oparciu o pewien zbiór treningowy $S = \{x_i, d_i\}_{i=1}^N$. Prezentowany model posiadał uniwersalne cechy pozwalające na realizację

podstawowych funkcji systemów uczących się, wśród których można wymienić kojarzenie, rozpoznawanie i klasyfikację wzorców. Okazuje się, że zaprezentowany system uczenia maszynowego można rozszerzyć do systemu Q-inspirowanego i zastosować do realizacji odwzorowań zdefiniowanych w przestrzeni liczb zespolonych $d = F(x)$, $d \in C^m, x \in C^n$ lub też przekształcających zbiory wektorów zdefiniowanych w przestrzeni liczb rzeczywistych na zbiory wektorów z przestrzeni liczb zespolonych $d = F(x)$, $d \in C^m, x \in R^n$. Możliwe są także przekształcenia odwrotne $d = F(x)$, $d \in R^m, x \in C^n$. Wymienione cechy badanego modelu pozwalają przykładowo na realizację przy jego pomocy dyskretnej transformacji Fouriera i odwrotnej transformacji Fouriera.

Autorzy: dr inż. Wiesław Ciłko, Uniwersytet Morski w Gdyni, Wydział Elektryczny, ul. Morska 81-87, 81-225 Gdynia, E-mail: w.citko@we.umg.edu.pl; dr hab. inż. Wiesław Sieńko, Uniwersytet Morski w Gdyni, Wydział Elektryczny, ul. Morska 81-87, 81-225 Gdynia, E-mail: w.sienko@we.umg.edu.pl.

LITERATURA

- [1] Nguyen N.H., Behrman E.C., Benchmarking Neural for Quantum Computations, *arXiv:1807.03253v3*, (2018).
- [2] Feng Q.L., Chao Chen, Qubit Neural Tree Networks with Applications in Nonlinear System Modelling, *IEEE Access*, vol.6, (2018).
- [3] Ciłko W., Sieńko W., Realizacja pamięci skojarzeniowej z zastosowaniem uczenia maszynowego, *Przegląd Elektrotechniczny*, nr 8, (2017), 77-80.
- [4] Ciłko W., Sieńko W., Zastosowaniem uczenia maszynowego do realizacji procesora analogowego, *Przegląd Elektrotechniczny*, nr 9, (2018), 56-58.
- [5] Ciłko W., Sieńko W., Realizowalność algorytmów kwantowych z zastosowaniem opartych na sieciach neuronowych modeli uczenia maszynowego, *Przegląd Elektrotechniczny*, nr 9, (2019), 146-149.