POLITECHNIKA ŚLĄSKA W GLIWICACH
WYDZIAŁ INŻYNIERII BIOMEDYCZNEJ

Projekt

Biocybernetyka
Uczenie sztucznych sieci neuronowych - przegląd metod

Daniel Imiołek

Zabrze, 4 stycznia 2018

Spis treści

1. Wprowadzenie

1.1 Sztuczna Sieć neuronowa - co to właściwie jest?

Definicją sztucznej sieci neuronowej jest zbiór prostych jednostek obliczeniowych przetwarzających dane, komunikujących się ze sobą i pracujących równolegle.

1.2 Typy Sieci Neuronowych

Wyróżniamy 3 typy sieci neuronowych

- Sieci Jednokierunkowe
- Sieci Rekurencyjne
- Samoorganizujące się mapy

2. Sztuczne Sieci Neuronowe

2.1 Metoda reguły delta

Reguła Delta została opracowana przez Widrowa i Hoffa, znalazła ona zastosowanie do uczenia elementów liniowych i nieliniowych. Reguła delta jest regułą uczenia z nauczycielem. Polega ona na tym, że każdy neuron po otrzymaniu na swoich wejściach określone sygnały (z wejść sieci albo od innych neuronów, stanowiących wcześniejsze piętra przetwarzania informacji) wyznacza swój sygnał wyjściowy wykorzystując posiadaną wiedzę w postaci wcześniej ustalonych wartości współczynników wzmocnienia (wag) wszystkich wejść oraz (ewentualnie) progu. Sposoby wyznaczania przez neurony wartości sygnałów wyjściowych na podstawie sygnałów wejściowych omówione zostały dokładniej w poprzednim rozdziale. Wartość sygnału wyjściowego, wyznaczonego przez neuron na danym kroku procesu uczenia porównywana jest z odpowiedzią wzorcową podaną przez nauczyciela w ciągu uczącym. Jeśli występuje rozbieżność - neuron wyznacza różnicę pomiędzy swoim sygnałem wyjściowym a tą wartością sygnału, która była by - według nauczyciela prawidłowa. Ta różnica oznaczana jest zwykle symbolem greckiej litery delta i stąd nazwa opisywanej metody.

Sygnał błędu wykorzystywany jest przez neuron do korygowania swoich współczynników wagowych:

- wagi zmieniane są tym silniej, im większy jest błąd
- wagi związane z tymi wejściami, na których występowały duże wartości sygnałów wejściowych zmieniane są bardziej niż wagi wejść, na których sygnał wejściowy był niewielki

Znając błąd popełniony przez neuron oraz jego wagi wejsciowe mozemy latwo przewidziec jak beda się zmieniac jego wagi.

$$y = \sum_{i=1}^{n} w_i x_i = \mathbf{W}^T \mathbf{X},$$

gdzie:

$$\mathbf{X} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix}, \qquad \mathbf{W} = \begin{bmatrix} w_1 \\ w_2 \\ w_3 \\ \vdots \\ w_n \end{bmatrix}$$

2.2 Metoda wstecznej propagacji błędów

Przez wiele lat nie znaleziono skutecznej metody uczenia sieci wielowarstwowych, dopiero w latach 80-tych zapropownowany został algorytm wstecznej propagacji błędów polegający na tym, że mając wyznaczony błąd $\delta_m^{(j)}$ występujący podczas realizacji j-tego kroku procesu uczenia w neuronie o numerze m można podawać ten błąd wstecz do wszystkich tych neuronów, których sygnały stanowiły wejścia dla m-tego neuronu.

Uczenie odbywa się przez minimalizację odpowiednio zdefiniowanej funkcji celu Q(W), przy czym wektor W reprezentuje wagi sieci poddawane optymalizacji. Najprostsza funkcja celu ma postać błędu średniokwadratowego. Zastosowanie różniczkowalnej funkcji aktywacji umożliwia minimalizację funkcji celu metodami gradientowymi.

Najprościej mówiąc celem tej metody jest zoptymalizowanie wag, aby sieć neuronowa mogła się nauczyć poprawnie mapować dowolne wejścia na wyjścia.

Schemat działania metody wstecznej propagacji błędów.

1 Wyznaczenie odpowiedzi neuronów warstwy wyjściowej oraz warstw ukrytych na zadany sygnał wejściowy.

2 Wyznaczenie błędu popełnianego przez neurony znajdujące się w warstwie wyjściowej i przesłanie go w kierunku warstwy wejściowej.

3 Adaptacja wag.

2.3 Sieć Hopfielda

`http://th-www.if.uj.edu.pl/~erichter/dydaktyka/Dydaktyka2012/SieciNN-2012/NN-wyklad`

`https://www.mimuw.edu.pl/~rlatkows/publications/latkowski1999sieci.pdf`

`Pseudoinwersja z instrukcji`

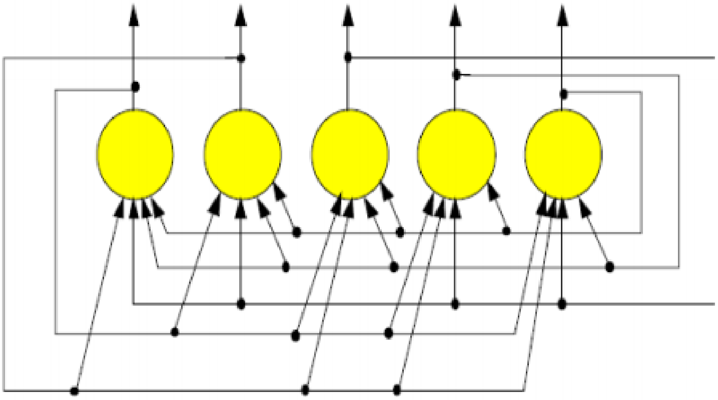
Sieć Hopfielda jest najbardziej znaną siecią, w której kierunek przepływu sygnałów jest odwrócony, posiada sprzężenia zwrotne typu każdy z każdym, jest prostym przykładem sieci rekurencyjnej i często jest nazywana autoasocjatorem, a w ramach tego sprzężenia każdy neuron jest połączony z jednym z wejść oraz z własnym wyjściem.

W sieciach hopfielda wykorzystujemy uczenie oparte na pseudoinwersji macierzy. Tak dobieramy wagi, aby uzyskać na wyjściu takie same wzorce jakie podajemy na wejściu.

$$WX = X \tag{2.1}$$

gdzie **W** to macierz wag o wymiarze n x n, a **X** to macierz wzorców o wymiarze n x p złożoną z p wektorów uczących.

Jednym z ważniejszych osiągnięć pracy Hopfielda jest pojęcie funkcji energetycznej w sieciach neuronowych. Najważniejszą własnością funkcji energetycznej jest to, że zawsze maleje lub pozostaje stała, gdy układ ewoluuje zgodnie z regułą



Rys. 2.1: Schemat sieci Hopfielda

$$Z_i^n = sign(Z_i^n + \frac{1}{N}\sum_{j=1}^{N}\sum_{m\neq n}^{p} Z_i^m Z_j^m Z_i^n)$$

W ten sposób wzorce leżą w minimach lokalnych powierzchni funkcji energetycznej.

Dla sieci neuronowych funkcja energetyczna istnieje, gdy wagi połączeń są symetryczne.

2.4 Sieci Kohonena

Tekst wyrównany do lewej.

Tekst wyśrodkowany.

Tekst wyrównany do prawej.

2.5 Podrozdział 1

2.6 Podrozdział 2

- **Pogrubiony tekst.**

- *Tekst pisany kursywą.*

1. punkt pierwszy

2. punkt drugi

2.6.1 Znaki Specjalne

Hasztag #
Backslash \ Dolar $

<table>
<tr><td>kom 11</td><td>kom 12</td></tr>
<tr><td colspan="2" align="center">kom 22 i 23</td></tr>
<tr><td colspan="2">kom 31 i kom 32</td></tr>
<tr><td>kom 41</td><td>kom 42</td></tr>
</table>

Tab. 2.1: tabela 1

odwołanie do tabeli (tab 2.1):

```
\begin{table} [h]
 \begin{tabular}{|l|c|p{7cm}|}
 \hline
 kom 11 & kom 12  \\
 \hline
 \hline
\multicolumn{2}{|c|}{kom 22 i 23}  \\
\hline
\multicolumn{2}{|l|}{kom 31 i kom 32}   \\
\hline
 kom 41 & kom 42  \\
 \hline
 \end{tabular}
 \centering
 \caption{tabela 1}\label{tab_1}
```

$$\left(\prod_{i=\widetilde{j}}^{\infty}[\log(i^{\xi})]^{M}\leqslant 0 \leftrightarrow \sum_{i=\widetilde{j}}^{\infty}\sqrt{\frac{i}{\widetilde{j}}} > \sqrt[p]{\widetilde{j}}\right) \Rightarrow \text{nic nie } \mathit{wynika} \tag{2.2}$$

Odwołanie do równania (2.2):

```
\begin{equation}\label{moje_równanie}
\left(\prod_{i=\widetilde{j}}^{\infty}[\log(i^{\xi})]^{M}\le0
 \leftrightarrow
  \sum_{i=\widetilde{j}}^{\infty}\sqrt{{i}\over{\widetilde{j}}} >
\sqrt[p]{\widetilde{j}}\right )\Rightarrow \textrm{nic nie}\ wynika
\end{equation}
```

Odwołanie do rysunku (rys 2.2):

Rys. 2.2: dwa rysunki jeden nad drugim

```
\begin{figure}
\raggedright
\includegraphics[scale=0.35]{rysunek1.jpg}
\includegraphics[scale=0.35]{rysunek1.jpg}
\caption{ dwa rysunki jeden nad drugim}\label{rysunek_1}

\end{figure}
```

3. Bibliografia

http://www.dbc.wroc.pl/Content/1908/Rusiecki_Algorytmy_PhD.pdf

https://platforma.polsl.pl/rib/pluginfile.php/2498/mod_resource/content/2/Laborator

http://www.neurosoft.edu.pl/media/pdf/tkwater/sztuczna_inteligencja/2_alg_ucz_ssn.p