

Met. techn. syst. w medyc. 2

Sprawozdanie

Temat: Komputerowe wspomaganie diagnozowania stanów ostrego brzucha z wykorzystaniem sztucznych sieci neuronowych

1. Cel projektu

Sprawdzić, jak jakość klasyfikacji (mierzona częstością poprawnej decyzji) zależy od liczby stosowanych cech. Dla sztucznych sieci neuronowych – sieć jednokierunkowa z 1 warstwą ukrytą dla 3 różnych liczby neuronów w warstwie ukrytej oraz dla uczenia metodą propagacji wstecznej z momentum i bez momentum.

2. Opis problemu medycznego jako zadania klasyfikacji

Dla danego problemu istnieje 8 klas, klas diagnozy, które są zdefiniowane następująco:

Klasy (diagnozy)	Ilość osobników danej klasy
1. Ostre zapalenie wyrostka robaczkowego	141
2. Zapalenie uchyłków jelit	17
3. Niedrożność mechaniczna jelit	29
4. Perforowany wrzód trawienny	28
5. Zapalenie woreczka żółciowego	55
6. Ostre zapalenie trzustki	32
7. Niecharakterystyczny ból brzucha	157
8. Inne przyczyny ostrego bólu brzucha	17

Każdy osobnik posiada 30 cech dyskretnych takich jak:

Płeć, Wiek, Lokalizacja bólu na początku zachorowania Lokalizacja bólu obecnie, Intensywność bólu, Czynniki nasilające ból, Czynniki przynoszące ulgę, Progresja bólu, Czas trwania bólu, Charakter bólu na początku zachorowania, Charakter bólu obecnie, Nudności i wymioty, Apetyt, Wypróżnienia, Oddawanie moczu, Poprzednie niestrawności, Żółtaczka w przeszłości, Poprzednie operacje brzuszne, Leki, Stan psychiczny, Skóra, Temperatura (pacha), Tętno, Ruchy oddechowe powłok brzusznych, Wzdęcia, Umiejscowienie bolesności uciskowej, Objaw Blumberga, Obrona mięśniowa, Wzmożone napięcie powłok brzusznych, Opory patologiczne, Objaw Murphy'ego.

Liczność populacji wynosi: 476 osobników.

3. Przedstawienie stosowanego algorytmu

- Oczyszczenie i przygotowanie danych

Z dostarczonego pliku Excel z danymi zostały usunięte puste kolumny, a nazwy kolumn zostały przesunięte do nagłówków kolumn

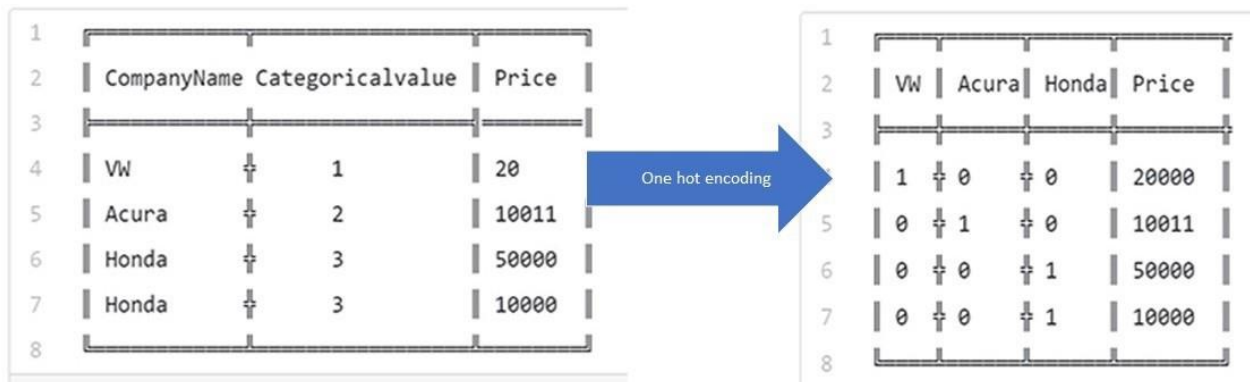
- Wyznaczyć ranking cech pod względem ich przydatności do klasyfikacji

W tym celu został obliczony współczynnik korelacji Pearsona, który uszeregował istotność cech następująco:

'Leki', 'Intensywność bólu', 'Czas trwania bólu', 'Stan psychiczny', 'Tętno', 'Objaw Murphy'ego', 'Umiejscowienie bolesności uciskowej', 'Wypróżnienia', 'Lokalizacja bólu obecnie', 'Wzmożone napięcie powłok brzusznych', 'Poprzednie operacje brzuszne', 'Charakter bólu na początku zachorowania', 'Lokalizacja bólu na początku zachorowania', 'Wiek', 'Płeć', 'Temperatura (pacha)', 'Opory patologiczne', 'Żółtaczka w przeszłości', 'Poprzednie niestrawności', 'Apetyt', 'Wzdęcia', 'Progresja bólu', 'Obrona mięśniowa', 'Oddawanie moczu', 'Czynniki nasilające ból', 'Skóra', 'Charakter bólu obecnie', 'Ruchy oddechowe powłok brzusznych', 'Czynniki przynoszące ulgę', 'Objaw Blumberga', 'Nudności i wymioty'

Selekcja cech wykazała że sieć uzyskuje najlepszą dokładność dla wszystkich cech.

- „One hot encoding” – kodowanie etykiet jako tablicę numeryczną, zastosowaliśmy tą metodę ze względu na to że dane nie są ze sobą skorelowane np. Lokalizacja bólu jest przedstawiana formie etykiet, które określają położenie. Przez co sieć ze względu na swoją strukturę i ilość warstw ukrytych, mogłaby nie zdążyć się wyuczyć. Wszystkie kolumny które zawierają więcej wartości niż 0 i 1 zostały poddane temu zabiegowi oprócz ostatniej kolumny z klasą diagnozy, co spowodowało około 30% wzrost dokładności.



- Przygotowanie danych do dwukrotnej walidacji krzyżowej
- Budowanie sieci neuronowej.

Zastosowana sieć neuronowa składa się z 3 warstw:

- z warstwy wejściowej (liczba neuronów wejściowych = 106, liczba neuronów wyjściowych = 106, funkcja aktywacji = 'relu'),
- warstwy ukrytej (dla warstwy ukrytej użytkownik deklaruje ilość neuronów, funkcja aktywacji = 'relu'),
- warstwy wyjściowej (liczba neuronów wyjściowych = 8, funkcja aktywacji = 'softmax').

Parametrem metody uczenia BP jest Stochastic Gradient Descent (SGD) – metoda spadku gradientu. Dla którego momentum jest ustawiane przez użytkownika w trakcie działania programu.

- Uczenie sieci – 30 epok
- Klasyfikacja – użyto klasyfikatora liniowego
- Stworzenie macierzy konfuzji

* 106 neuronów wejściowych w warstwie wejściowej wynika z zastosowania One hot encodingu dla najlepszego uzyskanego wyniku

* relu - $f(x) = \max(0, x)$

* softmax - Funkcja softmax lub inaczej znormalizowana funkcja wykładnicza, jest uogólnieniem funkcji sigmoidalnej, Transformuje ona K-wymiarowy wektor wartości rzeczywistych do K-wymiarowego wektora wartości w zakresie $(0; 1)$, którego wartości sumują się do 1. Funkcja ta jest często implementowana w ostatniej warstwie sieci wykorzystywanej do zadań klasyfikacji, gdy klasy wykluczają się wzajemnie. Wartość $f(x)_j$ reprezentuje prawdopodobieństwo przynależenia wartości wejściowej do klasy j .

4. Środowisko programistyczne

Do stworzenia projektu użyto Pythona, jupyter notebook, bibliotek matematycznych - scipy, numpy oraz biblioteki do tworzenia sieci neuronowych – keras.

5. Plan eksperymentu i wyniki

Plan:

- Uruchomić sieć dla 8, 49, 106 neuronów warstwie ukrytej bez momentum
- Uruchomić sieć dla 8, 49, 106 neuronów warstwie ukrytej z momentum
- Powtórzyć powyższe kroki po 5 razy dla różnych ilości cech

Wyniki:

- wyniki dla pierwszych 8 wyselekcjonowanych cech wg. rankingu

Liczba neuronów w warstwie ukrytej	Średnia dokładność sieci bez momentum	Średnia dokładność sieci z momentum
8	66.18%	32.98%
49	70.59%	32.14%
106	67.44%	48.32%

- wyniki dla pierwszy 15 wyselekcjonowanych cech wg. rankingu

Liczba neuronów w warstwie ukrytej	Średnia dokładność sieci bez momentum	Średnia dokładność sieci z momentum
8	78.15%	31.25%
49	85.71%	11.55%
106	88.24%	32.35%

- wyniki dla pierwszy 25 wyselekcjonowanych cech wg. rankingu

Liczba neuronów w warstwie ukrytej	Średnia dokładność sieci bez momentum	Średnia dokładność sieci z momentum
8	93.91%	32.98%
49	91.60%	29.62%
106	92.02%	31.30%

- wyniki dla pierwszy 30 wyselekcjonowanych cech wg. rankingu

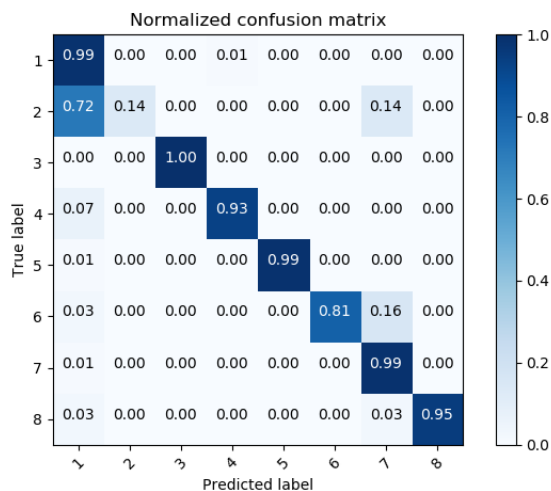
Liczba neuronów w warstwie ukrytej	Średnia dokładność sieci bez momentum	Średnia dokładność sieci z momentum
8	92.96%	42.44%
49	94.85%	32.98%
106	95.27%	25.84%

- wyniki dla wszystkich cech

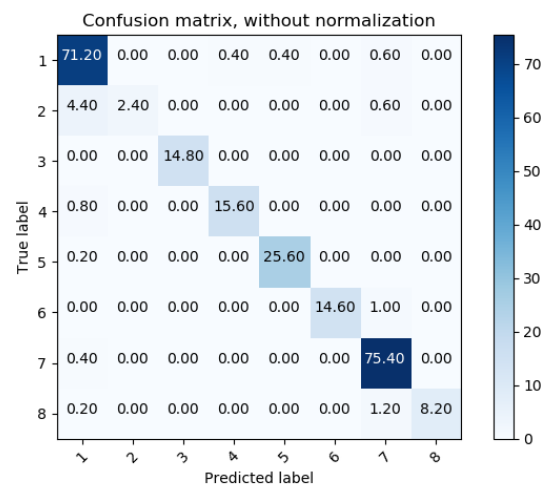
Liczba neuronów w warstwie ukrytej	Średnia dokładność sieci bez momentum	Średnia dokładność sieci z momentum
8	95.96%	29.15%
49	97.65%	27.69%
106	98.65%	23.15%

Uśrednione macierze konfuzji dla 106 liczb neuronów w warstwie ukrytej bez momentum.

Znormalizowana macierz
Wyniki przedstawione procentowo



Nie znormalizowana macierz
Wyniki przedstawione ilościowo



6.Wnioski

- Zauważono, że sieci z momentum uczą się o wiele szybciej dlatego w przypadku 30 epok widoczny jest efekt przeuczenia sieci i wyniki są o wiele gorsze.
- Wybrano wersję bez momentum i 30 epok dlatego że uzyskała lepsze wyniki.
- Ilość neuronów i cech ma wpływ na dokładność, w przypadku tego tematu okazało się, że czym więcej cech i czym więcej neuronów tym lepiej, dlatego też najlepsze wyniki otrzymano dla wszystkich cech.
- Dla mało licznych klas sieć często popełnia błędy.

Literatura

Dokumentacja keras'a - <https://keras.io/>

One hot encoding - <https://machinelearningmastery.com/how-to-one-hot-encode-sequence-data-in-python/>

Aurelie Geron, O'Reilly - Uczenie maszynowe z użyciem Scikit_Learn i TensorFlow