

Spis treści

Rozdział 1

Wstęp

1.1 Cel projektu

1.2 Motywacja

Rozdział 2

Problem medyczny

Wybrany przeze mnie problem medyczny dotyczy klasyfikacji stanów ostrego brzucha. Za ten stan odpowiedzialne mogą być różne choroby, które zawsze wymagają interwencji lekarza.

2.1 Opis chorób

Do klasyfikacji jest 8 chorób, zatem sieć neuronowa będzie miała za zadanie przypisać 1 z 8 klas. Są to:

1. Ostre zapalenie wyrostka robaczkowego
2. Zapalenie uchyłków jelit
3. Niedrożność mechaniczna jelit
4. Perforowany wrzód trawienny
5. Zapalenie woreczka żółciowego
6. Ostre zapalenie trzustki
7. Niecharakterystyczny ból brzucha
8. Inne przyczyny ostrego bólu brzucha

Histogram pokazuje, że rozkład klas jest nierównomierny. Na 476 obiektów aż 157 to 'Niecharakterystyczny ból brzucha' i 141 ma etykietę 'Ostre zapalenie wyrostka robaczkowego'. Czyli do 2 klas należy ponad 60% obiektów. Może to mieć negatywny wpływ na jakość klasyfikacji.

2.2 Opis cech

Dane do tego problemu zawierają 31 cech. Są to odpowiedzi z wywiadu medycznego i wyniki przeprowadzonych badań. Możliwe wartości parametrów przedstawione są poniżej. Jak widać wszystkie liczby są naturalne mniejsze niż 11, także normalizacja czy skalowanie danych nie jest konieczne.

- Ogólne

1. Płeć

- 1 - męska
- 2 - żeńska

2. Wiek

- 1 - poniżej 20 lat

- 2 - 20 - 30 lat
- 3 - 21 - 30 lat
- 4 - 31 - 40 lat
- 5 - 41 - 50 lat
- 6 - powyżej 50 lat

• Ból

3. Lokalizacja bólu na początku zachorowania

- 1 - prawa górna ćwiartka
- 2 - lewa górna ćwiartka
- 3 - górna połowa
- 4 - prawa połowa
- 5 - lewa połowa
- 6 - centralny kwadrat
- 7 - cały brzuch
- 8 - prawa dolna ćwiartka
- 9 - lewa dolna ćwiartka
- 10 - dolna połowa

4. Lokalizacja bólu obecnie

- 0 - brak bólu
- 1 - prawa górna ćwiartka
- 2 - lewa górna ćwiartka
- 3 - górna połowa
- 4 - prawa połowa
- 5 - lewa połowa
- 6 - centralny kwadrat
- 7 - cały brzuch
- 8 - prawa dolna ćwiartka
- 9 - lewa dolna ćwiartka
- 10 - dolna połowa

5. Intensywność bólu

- 0 - łagodny/brak
- 1 - umiarkowany
- 2 - silny

6. Czynniki nasilające ból

- 0 - brak czynników
- 1 - oddychanie
- 2 - kaszel
- 3 - ruchy ciała

7. Czynniki przynoszące ulgę

- 0 - brak czynników
- 1 - wymioty
- 2 - pozycja ciała

- 8. Progresja bólu
 - 1 - ustępujący
 - 2 - bez zmian
 - 3 - nasilający się
- 9. Czas trwania bólu
 - 1 - mniej niż 12 godzin
 - 2 - 12 - 24 godzin
 - 3 - 24 - 48 godzin
 - 4 - powyżej 48 godzin
- 10. Charakter bólu na początku zachorowania
 - 1 - przerywany
 - 2 - stały
 - 3 - kolkowy
- 11. Charakter bólu obecnie
 - 0 - brak bólu
 - 1 - przerywany
 - 2 - stały
 - 3 - kolkowy
- Inne objawy
 - 12. Nudności i wymioty
 - 0 - brak
 - 1 - nudności bez wymiotów
 - 2 - nudności z wymiotami
 - 13. Apetyt
 - 1 - zmniejszony
 - 2 - normalny
 - 3 - zwiększony
 - 14. Wypróżnienia
 - 1 - biegunki
 - 2 - prawidłowe
 - 3 - zaparcia
 - 15. Oddawanie moczu
 - 1 - normalne
 - 2 - dysuria
- Historia
 - 16. Poprzednie niestrawności
 - 0 - nie
 - 1 - tak
 - 17. Żółtaczka w przeszłości
 - 0 - nie

- 1 - tak
- 18. Poprzednie operacje brzuszne
 - 0 - nie
 - 1 - tak
- 19. Leki
 - 0 - nie
 - 1 - tak
- Ogólne badanie
 - 20. Stan psychiczny
 - 1 - pobudzony/cierpiący
 - 2 - prawidłowy
 - 3 - apatyczny
 - 21. Skóra
 - 1 - blada
 - 2 - prawidłowa
 - 3 - zaczerwieniona (twarz)
 - 22. Temperatura (pacha)
 - 1 - poniżej 36.5 stC
 - 2 - 36.5 - 37 stC
 - 3 - 37 - 37.5 stC
 - 4 - 37.5 - 38 stC
 - 5 - 38 - 39 stC
 - 6 - powyżej 39 stC
 - 23. Tętno
 - 1 - poniżej 60 /min
 - 2 - 60 - 70 /min
 - 3 - 70 - 80 /min
 - 4 - 80 - 90 /min
 - 5 - 90 - 100 /min
 - 6 - 100 - 110 /min
 - 7 - 110 - 120 /min
 - 8 - 120 - 130 /min
 - 9 - powyżej 130 /min
- Oglądanie brzucha
 - 24. Ruchy oddechowe powłok brzusznych
 - 1 - normalne
 - 2 - zniesione
 - 25. Wzdęcia
 - 0 - nie
 - 1 - tak

- Badania palpacyjne
 - 26. Umiejscowienie bolesności uciskowej
 - 0 - brak bólu
 - 1 - prawa górna ćwiartka
 - 2 - lewa górna ćwiartka
 - 3 - górna połowa
 - 4 - prawa połowa
 - 5 - lewa połowa
 - 6 - centralny kwadrat
 - 7 - cały brzuch
 - 8 - prawa dolna ćwiartka
 - 9 - lewa dolna ćwiartka
 - 10 - dolna połowa
 - 27. Objaw Blumberga
 - 0 - negatywny
 - 1 - pozytywny
 - 28. Obrona mięśniowa
 - 0 - nie
 - 1 - tak
 - 29. Wzmożone napięcie powłok brzusznych
 - 0 - nie
 - 1 - tak
 - 30. Opory patologiczne
 - 0 - nie
 - 1 - tak
 - 31. Objaw Murphy’ego
 - 0 - negatywny
 - 1 - pozytywny

2.3 Selekcja cech

Selekcja cech jest procesem wymaganym, gdy dane nie są dobrej jakości w wielu algorytmach uczenia maszynowego. Polega ona na wyborze podzbioru najlepszych cech według ustalonego kryterium. Analitycy danych przeprowadzają selekcję z następujących powodów:

- uproszczenie modelu, w celu ułatwienia interpretacji przez badaczy,
- skrócenie czasu treningu,
- zmniejszenie wymiarowości modelu,
- zwiększenie generalizacji poprzez uniknięcie zjawiska przeuczenia.

2.3.1 Test chi2

Metoda, którą wybrałem to test chi2. Jest to jedna z technik nieparametrycznych. Nada się bardzo dobrze do oceny istotności statystycznej cechy. Test ten polega na obliczeniu podanego poniżej wyrażenia dla każdej z cech i wybraniu takich, dla których wartość jest największa.

$$\chi^2 = \sum_{i=1}^n \frac{(O_i - E_i)^2}{E_i}$$

Gdzie:

- O_i - wartość mierzona,
- E_i - wartość oczekiwana,
- n - liczba obiektów.

Wartości testu dla wszystkich cech mają następujące wartości:

L.p.	Cecha	Wartość chi2
1	Charakter bólu obecnie	127.811
2	Czynniki przynoszące ulgę	87.453
3	Nudności i wymioty	84.633
4	Czas trwania bólu	84.273
5	Umiejscowienie bolesności uciskowej	77.456
6	Lokalizacja bólu obecnie	70.865
7	Czynniki nasilające ból	59.357
8	Tętno	58.152
9	Apetyt	54.489
10	Wypróżnienia	42.184
11	Charakter bólu na początku zachorowania	32.127
12	Lokalizacja bólu na początku zachorowania	31.430
13	Ruchy oddechowe powłok brzusznych	31.192
14	Progresja bólu	30.502
15	Objaw Blumberga	21.387
16	Wiek	21.228
17	Skóra	20.202
18	Intensywność bólu	18.438
19	Temperatura (pacha)	17.708
20	Stan psychiczny	15.930
21	Leki	15.554
22	Objaw Murphy'ego	13.666
23	Obrona mięśniowa	13.062
24	Oddawanie moczu	12.322
25	Wzmoczone napięcie powłok brzusznych	11.406
26	Wzdęcia	8.771
27	Opory patologiczne	8.504
28	Poprzednie operacje brzuszne	7.007
29	Płēć	6.195
30	Poprzednie niestrawności	4.470
31	Żółtaczka w przeszłości	0.590

Najlepszymi cechami są te, które mają wysoką wartość chi2. Zatem ograniczając liczbę cech, do klasyfikacji brane będą te z góry tabeli. Cechy o niskiej wartości, jak np. 'Żółtaczka w przeszłości', nie polepszają klasyfikacji, a mogą ją nawet pogorszyć.

Rozdział 3

Sieć neuronowa

3.1 Wprowadzenie

3.2 Neuron

3.2.1 Funkcja aktywacji

3.3 Model wielowarstwowy

Rozdział 4

Opis architektury aplikacji

4.1 Schemat warstwy

4.2 Schemat modelu

4.2.1 Proces uczenia

```
class Layer:
    def __init__(self, shape, activation='sigmoid'):
        ...

    def feedforward(self, x: np.ndarray) -> np.ndarray:
        ...

    def calc_delta(self, d: np.ndarray = None):
        ...

    def calc_gradient(self):
        ...

    def update_weights(self):
        ...
```

Powyższy fragment kodu przedstawia schemat klasy `Layer`. Jest to implementacja jednej warstwy w sieci neuronowej. Klasa zawiera w sobie tablicę, która jest składowa z wag połączeń do poprzedniej warstwy. Przy tworzeniu instancji można podać funkcję aktywacji (domyślnie jest to `sigmoid`).

Rozdział 5

Przeprowadzone badania

Rozdział 6

Podsumowanie

6.1 Dalsze możliwości rozwoju

6.2 Co mogłem zrobić lepiej

Tekst podsumowania

Spis rysunków

Spis tablic