Paweł Gradek

Systemy inteligentne Projekt

Klasyfikacja obiektów z metodą optymalizacji.

,,Problem plecakowy z klasyfikacją”- zastosowanie sieci neuronowej do klasyfikacji rodzynek oraz algorytmu tabu search do wybrania najlepszych rodzynek z wybranej klasy spełniających zadane ograniczenia.

Przedsiębiorstwo zajmujące się rozwojem nowych technologii w zakresie rolnictwa, bada jakość uprawianych odmian rodzynek. W celu określenia jakości odmiany Besni rodzynek poddaje się je procesowi badań. Aby badania miały jak największą wiarygodność wykorzystuje się jak największą liczbę rodzynek do przeprowadzenia badań wśród klasy Besni, które posiadają największy obwód. Używana technologia pozwala na ograniczoną wielkość powierzchni wszystkich rodzynek podanych procesowi badania.

W tym projekcie zaimplementowałem proces uczenia sieci neuronowej do klasyfikacji rodzynek. Skorzystałem ze zbioru dotyczącego rodzynek ,,raisin”. Zbiór danych ,,raisin” zawiera pomiary właściwości geometrycznych rodzynek dwóch różnych odmian rodzynek, po 450 zestawów atrybutów dla każdej odmiany.

Zadanie rozpoznawania polega na przyporządkowaniu obiektowi, charakteryzowanemu zestawem 7 cech, numeru kategorii ze zbioru {1,2}, w którym znajdują się numery klas.

*x* = []

Gdzie:

Badaną grupę stanowiły rodzynki należące do dwóch różnych odmian: Kecimen i Besni.

Prawidłowe numery klas dla zaobserwowanych zestawów atrybutów zostały podane przez ekspertów którzy udostępnili zbiór danych ,,raisin”.

Prawidłowe numery klas dla dwóch różnych odmian rodzynek:

Kecimen =1

Besni = 2

2. Źródłowy zbiór danych

Link do zbioru danych rodzynek: <https://archive.ics.uci.edu/ml/datasets/Raisin+Dataset>

3. Zaimplementowana struktura sieci

Liczba wejść sieci *k* = liczba atrybutów opisujących obiekty = 7

Liczba neuronów wyjściowych *r* = liczba klas = 2

Liczba warstw ukrytych i neuronów ukrytych może być dowolna, dlatego zdecydowałem się na 1 warstwę ukrytą i 5 neuronów ukrytych.

Liczba neuronów ukrytych *h* =5

Funkcje aktywacji w warstwach ukrytej i wyjściowej przyjąłem sigmoidalne.

Wzorcowe wyjścia dla zaobserwowanych zestawów atrybutów:

|  |  |
| --- | --- |
| Zestaw atrybutów określający odmianę rodzynek: | Wzorcowe wyjścia: |
| Kecimen | [1,0] |
| Besni | [0,1] |

Wartość 1 wpisuje na pozycji odpowiadającej numerowi klasy a na pozostałych pozycjach wpisuje 0.

4. Postać zbioru uczącego.

Zbiór uczący składa się z *N* = 700 zestawów, każdy zestaw zawiera 7 atrybutów i 1 wzorcowe wyjście.

[[,],

[,],

[,]]

Po klasyfikacji niektóre rodzynki z odmiany Besni są przeznaczone do badań. O tym które rodzynki odmiany Besni trafią do badań decyduje metoda optymalizacji. Mamy pojemnik do którego będziemy wkładać rodzynki przeznaczone do badań. Do tego pojemnika można włożyć rodzynki, których suma powierzchni nie przekracza *B* = 800 000 pikseli. Posiadamy N elementowy zbiór {}(liczba elementów tego zbioru będzie zależna od liczby rodzynek zaklasyfikowanych do klasy Besni), w którym każdy element ma określoną wartość (wielkość obwodu) oraz wielkość (pole powierzchni).

Formalnie problem może być zdefiniowany:

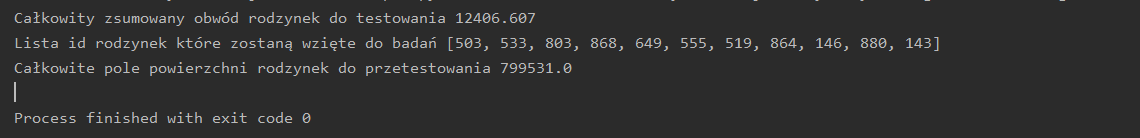
max

Przy założeniu: *B*

5. Rozwiązanie problemu podjęcia decyzji które rodzynki przeznaczyć do wykonania badań

W celu rozwiązania problemu zastosowałem algorytm tabu search, w którym przeszukuje przestrzeń, stworzoną ze wszystkich możliwych rozwiązań za pomocą sekwencji ruchów. Algorytm unika oscylacji wokół optimum lokalnego dzięki przechowywaniu informacji o sprawdzonych już rozwiązaniach w postaci listy tabu. Jeżeli przez 200 iteracji algorytm nie znajdzie lepszego rozwiązania od aktualnego to kończy działanie.

6. Podstawowe wyniki



Oznacza to, że do wykonania badań powinny być wzięte te rodzynki, których numer id znajduje się na liście. Mamy również podaną informacje jaki jest zsumowany obwód oraz całkowite pole powierzchni rodzynek, których powinniśmy wziąć do badań.

Bibliografia:

CINAR I., KOKLU M. and TASDEMIR S., (2020). “Classification of Raisin Grains Using Machine Vision and Artificial Intelligence Methods”, Gazi Journal of Engineering Sciences, vol. 6, no. 3, pp. 200-209, December, 2020, DOI: https://doi.org/10.30855/gmbd.2020.03.03

**DATASET:** https://www.muratkoklu.com/datasets/  
**Article Download (PDF):** https://dergipark.org.tr/tr/download/article-file/1227592