

DOKUMENTACJA KODU

Monika Kogut
Marta Kornaszewska
Katarzyna Kuzawska

Spis treści

czytaj_plik_excel	3
funkcja_bledu.....	3
funkcja_bledu_rozmyty	3
generuj_automat.....	3
generuj_csv	3
generuj_macierz	4
main.....	4
mapowanie_symboli	4
PSO	4
rozkład_normalny	4
stworz_rozmyty_zbior_uczacy	5
stworz_zbior_uczacy.....	5
symulacja_automatu.....	5
symulacja_automatu_rozmytego	6
TAIO2014	6
uzyj_zbioru_testowego	6
znajdz_symbol	6
znajdz_symbol_obcy.....	6

czytaj_plik_excel

Funkcja parsuje plik Excela, wydobywa informacje o liczbie kopii i tworzy wektor wszystkich symboli potrzebnych do testowania automatu.

Parametry:

- *sciezka_plik* - ścieżka do pliku Excel

funkcja_bledu

Wykorzystywana jest przez PSO w celu obliczenia błędu dla każdej iteracji. Macierz funkcji jest 3-wymiarowa, wygenerowana wcześniej przez funkcję 'generuj_automat'. Zbiór uczący wygenerowany wcześniej przez funkcję 'stworz_zbior_uczacy'. Liczba symboli i Liczba kopii potrzebna do funkcji *znajdz_symbol*.

Parametry:

- *wektor_z_macierzy3d* – macierz otrzymana z funkcji PSO w postaci jednowymiarowego wektora

funkcja_bledu_rozmyty

Wykorzystywana jest przez PSO w celu obliczenia błędu dla każdej iteracji. Macierz funkcji jest 3-wymiarowa, wygenerowana wcześniej przez funkcję 'generuj_automat'. Zbiór uczący wygenerowany wcześniej przez funkcję 'stworz_zbior_uczacy'. Liczba symboli i Liczba kopii potrzebna do funkcji *znajdz_symbol*.

Parametry:

- *wektor_z_macierzy3d* – macierz otrzymana z funkcji PSO w postaci jednowymiarowego wektora

generuj_automat

Losuje 3-wymiarową macierz funkcji przejścia dla automatów deterministycznego i niedeterministycznego.

Parametry:

- *liczba_symboli* – liczba klas automatu
- *dyskretyzacja* – podział zbioru wartości cech na równe przedziały

generuj_csv

Generuje plik csv z losowymi wartościami cech. Zakładamy, że mamy litery a, b, c, d... w zależności od liczby symboli. Cechy będą losowane z przedziału od 0 do *max_wartosc*.

Parametry:

- *l_symboli* – liczba klas w automacie
- *l_cech* – liczba cech każdego symbolu
- *min_wartosc* – minimalna wartość cechy *max_wartosc* – maksymalna wartość cechy
- *plik_wejsciowy* – plik do zapisania wygenerowanych wektorów symboli

generuj_macierz

Modyfikuje macierz zwracaną przez funkcję PSO w sposób zależny od automatu, dla deterministycznego wybiera maksymalną wartość w kolumnie i w jej miejsce wstawia 1, natomiast reszta wartości w danej kolumnie to 0; dla niedeterministycznego wybiera maksymalnie k największych wartości w kolumnie (k to podane przy wywołaniu ograniczenie automatu niedeterministycznego) i w ich miejsce wstawia 1, reszta to 0; dla automatu niedeterministycznego; dla rozmytego nie ma dyskretyzacji, elementy macierzy są losowe.

Parametry:

- *macierz* – macierz otrzymana z funkcji PSO

main

Zarządza działaniem całego programu, używa zmiennych wpisanych w wywołaniu TAIO2014.

mapowanie_symboli

Generuje wektor gdzie indeksom od 0 do liczba_symboli odpowiadają symbole z tablicy wejściowej wykorzystanej na koniec działania programu - gdy będziemy odczytywać wynik dla elementów testowych.

Parametry:

- *csvfile* – plik csv z danymi wektorami
- *liczba_symboli* – liczba klas automatu

PSO

Zewnętrzna funkcja PSO.

Parametry:

- *Objfunc* – handler do funkcji błędu
- *Nvars* – wymiar przestrzeni rozwiązań, w naszym przypadku jest to wymiar macierzy funkcji przejścia
- *Options* – opcje potrzebne do uruchomienia pso

rozklad_normalny

Wyznacza wartości rozkładu normalnego na podstawie danej średniej i wariancji.

Parametry:

- *srednia* – reprezentuje średnią wartość zaburzenia zbioru wartości, w naszym przypadku zawsze równa 0
- *wariancja* – zaburzenie zbioru wartości

stworz_rozmyty_zbior_uczacy

Generuje dane do zbioru uczącego automatu rozmytego, zwiększa ich ilość przez rozkład normalny, normalizuje i dzieli na przedziały. Funkcja zwraca gotowy zbiór uczący, w zależności od funkcji, zawierający elementy obce lub nie.

Parametry:

- *csvfile* – plik z którego czytamy wzorcowe wektory zawierające cechy symboli
- *liczba_symboli* – liczba klas automatu
- *liczba_cech* – liczba cech symboli
- *liczba_kopii* - liczba powtórzeń w klasie
- *dyskretyzacja* – podział zbioru wartości cech na równe podprzedziały
- *srednia* – reprezentuje średnią wartość zaburzenia zbioru wartości, w naszym przypadku zawsze równa 0
- *wariancja* – zaburzenie zbioru wartości
- *czy_plik* – zmienna mówiąca, czy zbiór uczący czytamy z pliku czy sami generujemy

stworz_zbior_uczacy

Generuje dane do zbioru uczącego, zwiększa ich liczbę przez rozkład normalny, normalizuje i dzieli na przedziały. Funkcja zwraca gotowy zbiór uczący, w zależności od funkcji, zawierający elementy obce lub nie.

Parametry:

- *csvfile* – plik z którego czytamy wzorcowe wektory zawierające cechy symboli
- *liczba_symboli* – liczba klas automatu
- *liczba_cech* – liczba cech symboli
- *liczba_kopii* - liczba powtórzeń w klasie
- *dyskretyzacja* – podział zbioru wartości cech na równe podprzedziały
- *srednia* – reprezentuje średnią wartość zaburzenia zbioru wartości, w naszym przypadku zawsze równa 0
- *wariancja* – zaburzenie zbioru wartości
- *czy_plik* – zmienna mówiąca, czy zbiór uczący czytamy z pliku czy sami generujemy

symulacja_automatu

Symulacja pracy automatu deterministycznego lub nondeterministycznego polegająca na mnożeniu macierzy (funkcji przejścia dla odpowiedniego symbolu) i kolejnych stanów automatu

Dla pierwszego mnożenie aktualny stan będzie losowym stanem ze wszystkich stanów pierwotnej konfiguracji automatu. Funkcja zwraca wynikowy stan, w jakim znajduje się automat po zakończeniu obliczeń.

Parametry:

- *wektor* – wektor reprezentujący symbol do rozpoznania przez automat
- *macierz_przejsc* – funkcja przejścia w automacie

symulacja_automatu_rozmytego

Symuluje pracę automatu, wykonuje mnożenie macierzy i wypisuje wynikowy stan automatu "wektor" to wektor ze zbioru uczącego w postaci np. [0.3 0.7 0.5 ...] w przeciwieństwie do poprzedniego etapu nie ma ściśle określonej przynależności do klasy, dlatego takie wartości Dla pierwszego mnożenia aktualny stan będzie wektorem o wartościach losowych z przedziału od 0 do 1 (jest to wektor stopni pewności). Funkcja zwraca wynikowy stan, w jakim znajduje się automat po zakończeniu obliczeń.

Parametry:

- *wektor* – wektor reprezentujący symbol do rozpoznania przez automat
- *macierz_przejsc* – funkcja przejścia w automacie

TAIO2014

Funkcja uruchamiająca program, przypisuje wartości odpowiednim zmiennym

uzyj_zbioru_testowego

Używa zbudowanego automatu do testowania symboli ze zbioru testowego(=wylosowanych wartości ze zbioru uczącego). Dla każdego wektora ze zbioru odbywa się symulacja automatu i odpowiednio zwiększa się błąd. Funkcja zwraca całkowity błąd będący stosunkiem błędnie rozpoznanych symboli do wszystkich symboli, a także wektor wynikowy zawierający rozpoznane symbole - używamy go do zapisania wyników w pliku Excela.

Parametry:

- *ilosc_elem* – ilość wektorów wchodzących w skład zbioru testowego
- *macierz_przejsc* – funkcja przejścia automatu
- *wektor_symboli* – wektor zmapowanych symboli do rozpoznania

znajdz_symbol

Znajduje indeks symbolu, jaki reprezentowany jest przez wektor o indeksie "indeks" ze zbioru uczącego. Zwraca indeks, który jest wykorzystywany w wektorze zmapowanych symboli lub -1 w przypadku, gdy symbol jest obcy.

Parametry:

- *indeks* – indeks wektora reprezentującego symbol w zbiorze uczącym
- *liczba_symboli* – liczba klas automatu
- *liczba_kopii* – ilość powtórzeń w klasie

znajdz_symbol_obcy

Sprawdza, czy symbol reprezentowany przez wektor *wektor_sym* został odrzucony przez automat. Funkcja zwraca 1, jeśli symbol został odrzucony lub -1 w przeciwnym przypadku.

Parametry:

- *wektor_sym* – wektor reprezentujący dany symbol