

Sprawozdanie

Ćwiczenie nr. 2
"Miara odległości Euclidesa"

Wykonane przez: Jarosław Jędruszcak

1. Wprowadzenie

Na podstawie poprzedniego ćwiczenia należało wyznaczyć odległości między wybranym dokumentem a wszystkimi co pozwoliło nam na określenie podobieństw między dokumentami metodą Euclidesową.

2. Miara odległości Euclidesowa

Jest to odległość geometryczna w przestrzeni wielowymiarowej. Obliczana wzorem:

$$\text{odległość}(x,y) = \{\sum_i (x_i - y_i)^2\}^{1/2}$$

gdzie:

- x oraz y = dokumenty porównywane do siebie
- i = wszystkie termsy

Wzór ten wyznaczy sumę różnicy termów między dwoma dokumentami podniesioną do kwadratu a następnie po spierwiaztkowaniu sumy uzyskujemy odległość między dokumentem „x” a „y”.

Wyniki które powstają następnie się porównuje w sposób następujący: dokument który ma najniższa wartość odległości jest najbardziej podobny do porównywanego dokumentu np.:

D2 - D1	9.53939...
D2 - D2	0.0
D2 - D3	6.70820...

Powyżej widoczne są trzy porównania między dokumentem 2 a pozostałymi. Dokument który jest najbardziej podobny do dokumentu drugiego jest dokument trzeci ponieważ ma mniejszą wartość odległości od porównywania z pierwszym dokumentem. Wartość 0 jest pomijana ponieważ wykazuje ona tylko że dokument został porównany ze sobą co daje wynik odległości 0.

3. Uzyskana macierz porównań

Który dokument chcesz porównać						TFM > Euq	Reset			
	drużyny	piłka	boisko	sport	gra	zawodnicy	punkty	zdrowie	Suma	Euclides
D2 - D1	49.0	1.0	0.0	4.0	36.0	1.0	0.0	0.0	91	9.53939...
D2 - D2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0	0.0
D2 - D3	0.0	0.0	1.0	36.0	0.0	4.0	4.0	0.0	45	6.70820...
D2 - D4	0.0	0.0	1.0	1.0	100.0	0.0	16.0	0.0	118	10.8627...
D2 - D5	0.0	0.0	9.0	36.0	9.0	4.0	0.0	0.0	58	7.61577...
D2 - D6	0.0	0.0	0.0	16.0	100.0	4.0	0.0	0.0	120	10.9544...
D2 - D7	0.0	1.0	1.0	16.0	4.0	1.0	0.0	0.0	23	4.79583...
D2 - D8	0.0	0.0	0.0	0.0	49.0	0.0	1.0	0.0	50	7.07106...
D2 - D9	0.0	0.0	0.0	64.0	100.0	1.0	0.0	0.0	165	12.8452...
D2 - D10	0.0	0.0	0.0	121.0	144.0	4.0	0.0	49.0	318	17.8325...

Wynikiem jest ostatnia kolumna podpisana słowem „Euclides”

4. Fragment kodu odpowiedzialny za przedstawienie tabeli:

```
public static void obliczanieEuq() {  
  
    int indexDok = Integer.parseInt(EuqTargetTextF.getText()); //Zczytanie wartości która reprezentuje index dokumentu porównywanego z innymi  
    int sumaTermow = 0;  
  
    for (int i = 0; i < LdwTSM; i++) //For dla wierszy  
        for (int ii = 0; ii < (int)IleRoznychTermow+3; ii++) //For dla kolumn  
        {  
            if (ii == 0 || i == 0) { //Wiersz 0 kolumna 0 ma pozostać pusta  
                if (ii == 0) {  
                    TEuq.setValueAt("D" + indexDok + " - D" + i, i, ii); //1 kolumna zawiera nazwy porównywanych dokumentów tj. D1 - D9  
                }  
                if (i == 0) {  
                    if (ii < (int)IleRoznychTermow+1) //Pierwszy wiersz zawiera nagłówki  
                        TEuq.setValueAt(TSM.getValueAt(i, ii), i, ii); //Nagłówki termow  
                    else //reszta nagłówek jest specjalna  
                        if (ii == (int)IleRoznychTermow+1)  
                            TEuq.setValueAt("Suma", i, ii);  
                        else  
                            if (ii == (int)IleRoznychTermow+2)  
                                TEuq.setValueAt("Euqlides", i, ii);  
                }  
            }  
            else  
            {  
                if (ii < (int)IleRoznychTermow+1) //Wypelnianie wartosciami liczbowymi  
                {  
                    //Wyniesienie czech nazwy tj. różnica między dokumentami podniesiona do kwadratu  
                    Double wynik = Math.pow((Double.parseDouble((String) TSM.getValueAt(indexDok,ii)))-Double.parseDouble((String) TSM.getVal  
                    TEuq.setValueAt(Double.toString(wynik),i,ii);  
                    sumaTermow += wynik;  
                }  
                else  
                {  
                    if (ii == (int)IleRoznychTermow+1) // Suma wszystkich wyników wiersza  
                        TEuq.setValueAt(Integer.toString(sumaTermow), i, ii);  
                    else  
                    {  
                        if (ii == (int)IleRoznychTermow+2)  
                        {  
                            // Pierwiastek sumy  
                            TEuq.setValueAt(Double.toString(Math.sqrt(sumaTermow)), i, ii);  
                            sumaTermow = 0;  
                        }  
                    }  
                }  
            }  
        }  
    }  
}
```