

Raport Colocviu

Lazaroiu Teodora-Bianca, grupa 241

Exercitiul 2

Prin metoda celor mai apropiati vecini se obtine labelul prin calcularea similaritatii dintre datele de test si toate datele de train. Se vor alege primele k similaritati, unde am ales $k = 5$ in cazul de fata. Metoda KNN clasifica in functie de cat de apropiata este imaginea de test de imaginile de train insa, in acest caz, gradul de apropiere este reprezentat de similaritatea calculata prin functia de la punctul 1

```
In [ ]: diferente = np.zeros(len(train_data))

        for i in range(len(train_data)):
            diferente[i] = similaritate(train_data[i], test_data, 8)

        # pastrez cei mai apropiati K vecini

        # indicii elementelor sortate dupa suma
        indici = np.argsort(diferente)
        k = 5

        # returnez labelul cel mai frecvent printre vecini
        labels = []
        for i in range(k):
            labels.append(train_labels[indici[i]])

        label = max(labels, key = labels.count)

        return label
```

La final este generat un fisier de tip text unde sunt scrise pe fiecare linie etichetele corespunzatoare imaginilor de test calculate prin metoda celor mai apropiati vecini. Fisierul are 323 de linii si cuprinde valori de 1 sau -1.