

Praktyczne zastosowanie metod SI – opis projektu dla grupy A

Temat: regularyzowany autoenkoder

Proszę o przygotowanie w języku Python aplikacji posiadającej następujące funkcje:

1. Obsługa danych:
 - a. Wczytanie danych z pliku (m. in. MNIST).
 - b. Wybór jednego z wcześniej ustalonych zbiorów wczytanych z pakietu scikit-learn (m. in. Olivetti Faces, digits)
2. Znalezienie składowych głównych danych przy wykorzystaniu regularyzowanego autoenkodera.
3. Zapis wyuczonego modelu do pliku / wczytanie wyuczonego modelu.
4. Wyświetlenie pierwszych n składowych głównych i opcja zapisu do pliku.
5. Wczytanie rekordu lub zbioru rekordów testowych. Rekonstrukcja danych, wyświetlenie rekonstrukcji z opcją zapisu do pliku, średniego błędu rekonstrukcji, średniej kwadratów błędów rekonstrukcji.

Autoenkoder – sieć neuronowa wykonująca zadanie PCA (liczba wejść = liczba wyjść) z nieliniową funkcją aktywacji. Na początek można przetestować z liniową i zobaczyć czy uzyskujemy pożądany efekt.

Regularyzacja – technika polegająca na wyzerowaniu części wag sieci neuronowej. Pozwala na zmniejszenie rozmiaru modelu oraz (w przypadku PCA) uzyskanie rzadkich składowych głównych.