LABORATORIUM 6

Zadanie implementacja PCA dla danych wielowymiarowych

1.DANE WEJŚCIOWE, PRZYGOTOWANIE DANYCH DO ALGORYTMU

Obraz zawierający tekst, Czcionka, zrzut ekranu

Zawartość wygenerowana przez sztuczną inteligencję może być niepoprawna.

Na początek implementacji ustaliłem jako dane wejściowe:

X – baza danych,

n – ilość wymiarów.

Następnie przygotowałem dane do dalszych kroków PCA w postaci obliczenia średniej i centrowania danych, aby każda cecha miała średnią zero.

2.MACIERZ KOWARIANCJI, WARTOŚCI I WEKTORY WŁASNE

Obraz zawierający tekst, Czcionka, zrzut ekranu, typografia

Zawartość wygenerowana przez sztuczną inteligencję może być niepoprawna.

Kolejnym krokiem było stworzenie macierzy kowariancji. W dalszej części za pomocą linalg.eigh obliczamy wartości i wektory własne, które opisują kierunki w przestrzeni danych gdzie rozkład danych ma największą wariancje (zróżnicowanie). W kolejnym etapie te wartości sortujemy w kolejności malejącej.

3.ZREDUKOWANIE WYMIARÓW

Obraz zawierający tekst, Czcionka, zrzut ekranu, linia

Zawartość wygenerowana przez sztuczną inteligencję może być niepoprawna.

W końcowym etapie wybieramy n wektorów własnych (w kolejności od najważniejszego). Rzutujemy dane przekształcając na nową wymiarowość.

Obraz zawierający tekst, zrzut ekranu, oprogramowanie, Czcionka

Zawartość wygenerowana przez sztuczną inteligencję może być niepoprawna.CAŁY KOD

Zadanie1. Implementacja PCA i funkcji dwuwymiarowej wizualizacji przestrzeni cech

1. wygenerować w sposób losowy zbiór 200 obiektów dwuwymiarowych za pomocą funkcji z numpy dot i rand lub randn

KOD NAPISANY

Obraz zawierający tekst, Czcionka, zrzut ekranu, Grafika

Zawartość wygenerowana przez sztuczną inteligencję może być niepoprawna.

1. zwizualizować obiekty na pomocą funkcji matplotlib, np. scatter

KOD NAPISANY

1. dokonać redukcji do jednego wymiaru za pomocą własnej funkcji wiPCA i zwizualizować wektory własne oraz rzut wygenerowanych obiektów na pierwszą składową, w sposób podobny do Rysunku 1

KOD NAPISANY

Obraz zawierający tekst, zrzut ekranu, Czcionka

Zawartość wygenerowana przez sztuczną inteligencję może być niepoprawna.

Uruchamiamy nasz algorytm z prośba o redukcje do 1 wymiaru. Na podstawie nowych danych rekonstruujemy dane. Wcześniej musimy wybrać główne wektory, ponieważ algorytm zwraca wszystkie wektory. Wizualizujemy zrekonstruowaną macierz.

Dodatkowo dla wizualnych efektów zmieniamy kierunki wektorów jeśli jest taka potrzeba, ponieważ kierunki wektorów własnych są niejednoznaczne znakowo co oznacza, że znak minus i plus daje ten sam kierunek (nie zmienia to matematycznego działania algorytmu).Zmiana kierunków wektorów będzie zastosowana również w późniejszych zadaniach w celu lepszego zobrazowania przykładowych rysunków.

W kolejnym etapie rysujemy za pomocą funkcji plt.quiver strzałki pokazujące kierunki wektorów.

Obraz zawierający Wykres, diagram, linia

Zawartość wygenerowana przez sztuczną inteligencję może być niepoprawna.REZULTAT

ZADANIE 2

Testowanie PCA na zbiorze iris

1. Wczytać zbiór iris (sposób j.w.)
2. dokonać redukcji wymiarowości wszystkich obiektów w zbiorze do 2 najbardziej znaczących wymiarów za pomocą opracowanej funkcji wiPCA
3. Zwizualizować elementy zbioru w przestrzeni cech z oznaczonymi klasami, np. za pomocą kolorów, etykiet lub symboli (\*, x, +, .)

KOD Z WYKORZYSTANIEM WŁASNEJ FUNCKJI

Obraz zawierający tekst, zrzut ekranu, oprogramowanie, Oprogramowanie multimedialne

Zawartość wygenerowana przez sztuczną inteligencję może być niepoprawna.

KOD Z WYKORZYSTANIEM FUNCKJI SKLEARN

Obraz zawierający tekst, zrzut ekranu, oprogramowanie, Czcionka

Zawartość wygenerowana przez sztuczną inteligencję może być niepoprawna.

REZULTATY

Obraz zawierający tekst, zrzut ekranu, diagram

Zawartość wygenerowana przez sztuczną inteligencję może być niepoprawna.Obraz zawierający tekst, zrzut ekranu, diagram

Zawartość wygenerowana przez sztuczną inteligencję może być niepoprawna.

Zadanie 3. Testowanie PCA na zbiorze digits

1. Wczytać zbiór digits (load digits)
2. dokonać redukcji wymiarowości wszystkich obiektów w zbiorze do 2 najbardziej znaczących wymiarów za pomocą opracowanej funkcji wiPCA

KOD NAPISANY

Obraz zawierający tekst, zrzut ekranu, Czcionka

Zawartość wygenerowana przez sztuczną inteligencję może być niepoprawna.

1. Pokazać krzywą wariancji dla rosnącej liczby składowych głównych (tak jak na Rysunku 3)

KOD NAPISANY

Obraz zawierający tekst, zrzut ekranu, Czcionka

Zawartość wygenerowana przez sztuczną inteligencję może być niepoprawna.

Obraz zawierający tekst, linia, diagram, zrzut ekranu

Zawartość wygenerowana przez sztuczną inteligencję może być niepoprawna.Funkcja przyjmuje dane, posortowane wektory oraz ilość elementów. Tworzymy macierz zer następnie obliczamy wariancje dla rosnącej liczby składowych dzieląc sumę kolejnych składowych przez sumę wszystkich. Na koniec zostaje zwracany wektor tych wariancji i pokazany na wykresie.

REZULTAT

1. Zwizualizować elementy zbioru w przestrzeni cech z oznaczonymi klasami (podobnie do Rysunku 4, funkcja scatter)

KOD NAPISANY



Zwizualizowałem dla naszego algorytmu wiPCA jak i dla gotowego z biblioteki sklearn.

Obraz zawierający diagram, Wielobarwność

Zawartość wygenerowana przez sztuczną inteligencję może być niepoprawna.REZULTAT

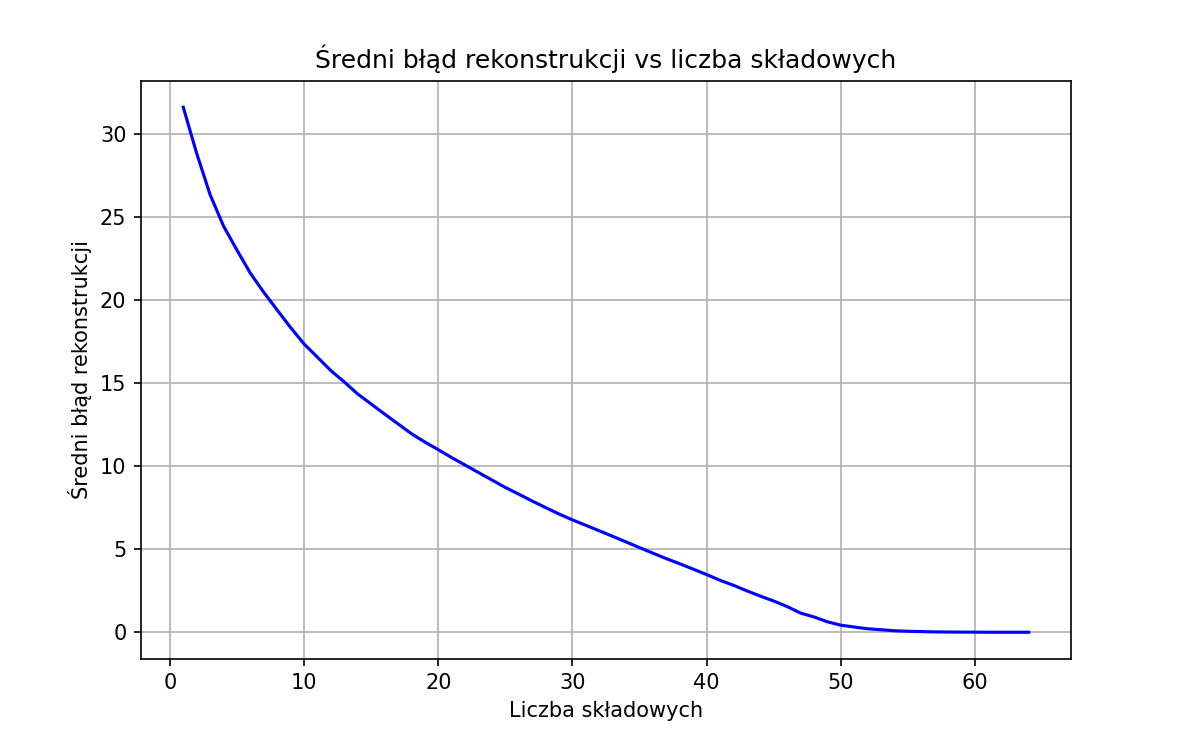
1. Wykonać eksperyment polegający na ocenie średniego błędu rekonstrukcji dla całego zbioru dla kolejno zwiększającej się liczby składowych głównych (można to zrobić za pomocą obliczania odległości dla wszystkich obiektów w bazie od ich zrekonstruowanych postaci - funkcja z Laboratorium nr 6) - przykładowy przebieg zmienności odległości na Rysunku 5. Zadanie to wymaga napisania funkcji obliczającej transformatę odwrotną do PCA, zwracającą obiekt(-y) o wymiarowości zgodnej z obiektem(-ami)

Obraz zawierający tekst, zrzut ekranu, Czcionka

Zawartość wygenerowana przez sztuczną inteligencję może być niepoprawna.KOD NAPISANY

Rekonstruujemy nowe dane a następnie obliczamy różnice miedzy nowo zrekonstruowanymi danymi a starymi. Na koniec wyświetlamy wizualnie listę z błędami.

REZULTAT



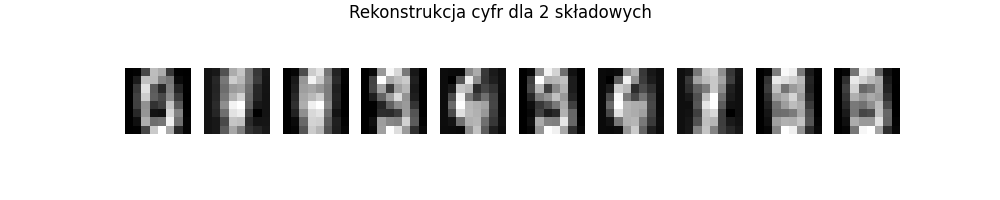
ZADANIE Dla ambitnych: pokazać zrekonstruowane cyfry dla 2, 4, 10 i 50 składowych głównych.

KOD NAPISANY

Obraz zawierający tekst, zrzut ekranu, Czcionka

Zawartość wygenerowana przez sztuczną inteligencję może być niepoprawna.

Wyświetlamy pierwsze 10 obrazków. Rekonstruujemy dane za pomocą wcześniej zrobionej funkcji odwrotna\_PCA. Po wyświetleniu rezultatu widać, że czym więcej składowych tym obraz jest bardziej wyraźny i nasycony. Rezultaty są na następnej stronie.

REZULTATY

