

TASK 1:

Obținem dimensiunile matricei `photo` și le atribuim variabilelor `m` și `n`.

Inițializăm matricea finală `new_X` cu dimensiunile `m x n`, inițializată cu elemente zero.

Convertim matricea `photo` la tipul de date `double` pentru a efectua calculele necesare.

Aplicăm algoritmul SVD (descompunerea în valori singulare) asupra matricei `photo`. Obținem trei matrici: matricea de vectori singolari stânga `U`, matricea diagonală a valorilor singulare `S` și matricea de vectori singolari dreapta `V`.

Reducem dimensiunea matricelor `U`, `S` și `V` pentru a obține matrici reduse `U_reduced`, `S_reduced` și `V_reduced`, păstrând doar primele `k` coloane.

Calculăm noua matrice `new_X` ca produsul matricilor `U_reduced`, `S_reduced` și `V_reduced'` (transpusa matricei `V_reduced`), reprezentând aproximarea matricii inițiale `photo`.

Convertim matricea `new_X` la tipul de date `uint8` pentru a obține o imagine validă (valori întregi în intervalul 0-255).

Returnăm matricea `new_X` ca rezultat al funcției.

TASK 2:

Obținem dimensiunile matricei `photo` și le atribuim variabilelor `m` și `n`.

Inițializăm matricea finală `new_X` cu dimensiunile `m x n`, inițializată cu elemente zero.

Convertim matricea `photo` la tipul de date `double` pentru a efectua calculele necesare.

Normalizăm matricea inițială `photo` prin scăderea mediei fiecărui rând. Calculăm media fiecărui rând și o atribuim variabilei `mu`. Scădem această medie din fiecare rând al matricei `photo`.

Construim matricea `Z` prin transpunerea matricei `photo` normalizate.

Aplicăm descompunerea în valori singulare (SVD) asupra matricei `Z`. Obținem trei matrici: matricea de vectori singolari stânga `U`, matricea diagonală a valorilor singulare `S` și matricea de vectori singolari dreapta `V`.

Construim matricea `W` din primele `pcs` coloane ale matricei `V`.

Calculăm matricea `Y` prin înmulțirea transpusa matricei `W` cu matricea `photo` normalizată.

Aproximăm matricea inițială `photo` utilizând matricele `W` și `Y`, adăugând înapoi media fiecărui rând `mu`.

Convertim matricea `new_X` la tipul de date `uint8` pentru a obține o imagine validă (valori întregi în intervalul 0-255).

Returnăm matricea `new_X` ca rezultat al funcției.

TASK 3:

Obținem dimensiunile matricei `photo` și le atribuim variabilelor `m` și `n`.

Inițializăm matricea finală `new_X` cu dimensiunile `m x n`, inițializată cu elemente zero.

Convertim matricea `photo` la tipul de date `double` pentru a efectua calculele necesare.

Calculăm media fiecărui rând al matricei `photo` și o atribuim variabilei `mu`.

Normalizăm matricea inițială `photo` prin scăderea mediei fiecărui rând din fiecare element corespunzător.

Calculăm matricea de covarianță a matricei `photo` normalizate utilizând formula covarianței sample-ului.

Obținem vectorii și valorile proprii ale matricei de covarianță folosind funcția `eig`.

Ordonăm valorile proprii în ordine descrescătoare și reordonăm corespunzător vectorii proprii în matricea `V`.

Păstrăm doar primele `pcs` coloane din matricea `V`, obținând astfel matricea de transformare `W`.

Construim matricea `Y` prin înmulțirea transpusa matricei `W` cu matricea `photo` normalizată.

Calculăm noua matrice `new_X` ca produsul matricilor `W` și `Y`, reprezentând aproximarea matricii inițiale `photo`.

Adunăm înapoi media fiecărui rând `mu` la matricea `new_X`.

Convertim matricea `new_X` la tipul de date `uint8` pentru a obține o imagine validă (valori întregi în intervalul 0-255).

Returnăm matricea `new_X` ca rezultat al funcției.

TASK 4:

-PREPARE_DATA:

Se definește variabila `n` cu valoarea 784, reprezentând numărul de pixeli ai unei imagini din setul de date MNIST.

Se inițializează matricea `train_mat` cu dimensiunea `no_train_images x n`, inițializată cu elemente zero. Ace

asta va stoca imaginile de antrenament.

Se inițializează vectorul `train_val` cu dimensiunea $1 \times \text{no_train_images}$, inițializat cu elemente zero. Acesta va stoca etichetele corespunzătoare imaginilor de antrenament.

Se încarcă datele din fișierul specificat în variabila `data`, utilizând funcția `load`.

Se extrag imaginile de antrenament din datele încărcate și se atribuie matricei `train_mat`. Imaginile sunt preluate din câmpul `trainX` al variabilei `data`, limitându-se la primele `no_train_images` rânduri.

Se extrag etichetele corespunzătoare imaginilor de antrenament și se atribuie vectorului `train_val`. Etichetele sunt preluate din câmpul `trainY` al variabilei `data`, limitându-se la primele `no_train_images` elemente.

Se returnează matricea `train_mat` și vectorul `train_val` ca rezultate ale funcției.

-VISUALISE_IMAGE:

Se inițializează matricea finală `im` cu dimensiunea 28×28 , inițializată cu elemente zero. Aceasta va reprezenta imaginea vizualizată.

Se citește rândul cu numărul specificat din matricea de antrenament `train_mat` și se atribuie variabilei `image_row`.

Se reorganizează rândul citit într-o matrice de dimensiune 28×28 utilizând funcția `reshape`, apoi se transpune matricea rezultată. Astfel, obținem matricea `im` care reprezintă imaginea vizualizată.

Se convertește matricea `im` la tipul de date `uint8` pentru a o transforma într-o imagine validă (valori întregi în intervalul 0-255).

Se afișează imaginea utilizând funcția `imshow`.

Funcția se încheie.

-MAGIC_WITH_PCA:

Se obțin dimensiunile matricei de antrenament `train_mat` și se atribuie variabilelor `m` și `n`.

Se inițializează matricea `train` cu dimensiunea $m \times n$, inițializată cu elemente zero. Aceasta va stoca rezultatul final al transformării PCA.

Se inițializează vectorul `miu` cu dimensiunea $1 \times n$, inițializat cu elemente zero. Acesta va reprezenta media a coloanelor matricei de antrenament.

Se inițializează matricea `Y` cu dimensiunea $m \times \text{pcs}$, inițializată cu elemente zero. Aceasta va reprezenta matricea rezultată după schimbarea bazei matricei de antrenament.

Se inițializează matricea `Vk` cu dimensiunea $n \times \text{pcs}$, inițializată cu elemente zero. Aceasta va reprezenta matricea de vectori proprii corespunzători componentelor principale.

Se convertește matricea de antrenament `train_mat` la tipul de date `double`.

Se calculează media fiecărei coloane a matricei `train_mat` utilizând funcția `mean` și se atribuie vectorului `miu`.

Se scade media `miu` din matricea de antrenament `train_mat` pentru a o centra în jurul originii.

Se calculează matricea de covarianță `cov_mat` a matricei de antrenament `train_mat` utilizând funcția `cov`.

Se calculează vectorii și valorile proprii ale matricei de covarianță `cov_mat` utilizând funcția `eig`.

Se sortează valorile proprii în ordine descrescătoare și se creează matricea `Vk` care conține vectorii proprii corespunzători sortati ca coloane.

Se păstrează doar primele `pcs` coloane din matricea `Vk`.

Se creează matricea `Y` prin schimbarea bazei matricei de antrenament utilizând matricea `Vk`.

Se calculează matricea `train` ca produsul matricial dintre matricele `Y` și `Vk'`, reprezentând astfel o aproximare a matricei de antrenament inițiale.

Funcția returnează matricea `train`, media `miu`, matricea `Y` și matricea `Vk` ca rezultate ale funcției.

-PREPARE_PHOTO:

Inițializăm vectorul final `sir` cu dimensiunea 1×784 , inițializat cu elemente zero. Acesta va stoca rezultatul final, adică vectorul de pixeli al imaginii pregătite.

Inversăm pixelii imaginii `im` prin scăderea valorii fiecărui pixel din 255. Această operație inversează culorile imaginii.

Transpunem imaginea `im` pentru a ne asigura că este în formatul corect pentru `reshape`.

Realizăm operația de `reshape`, transformând matricea într-un vector de pixeli. În acest caz, transpunerea ne asigură că valorile pixelilor vor fi plasate corect în vector.

Atribuim imaginea `reshaped` vectorului final `sir`.

Func ia returnează vectorul sir ca rezultat al func iei.

-KNN:

Ini ializăm variabila prediction cu valoarea -1. Aceasta va reprezenta eticheta prezisă pentru datele de test are.

Ini ializăm matricea de distan e distance cu dimensiunea $m \times 1$, unde m reprezintă numărul de exemple de antrenament.

Calculăm distan a Euclidiană între fiecare rând din matricea Y i vectorul de test test. Aceasta se realizează prin norma diferen ei dintre rândurile respective.

Sortăm distan ele în ordine crescătoare i păstrăm primele k valori. Sortarea este realizată folosind func ia sort.

Ob inem etichetele corespunzătoare celor k vecini cei mai apropia i i le atribuim variabilei `k_nearest_labels`.

Calculăm eticheta prezisă (prediction) ca mediană a etichetelor celor k vecini cei mai apropia i. Aceasta se realizează folosind func ia median.

Func ia returnează eticheta prezisă (prediction) ca rezultat al func iei.

-CLASSIFYIMAGE:

Ini ializăm variabila prediction cu valoarea -1. Aceasta va reprezenta eticheta prezisă pentru imaginea de testare.

Convertim imaginea `im` la tipul `double`.

Aplicăm func ia `magic_with_pca` asupra datelor de antrenament (`train_mat`), pentru a ob ine matricele `train` (antrenamentul transformat prin PCA), `miu` (media coloanelor din `train_mat`), Y (caracteristicile transformate prin PCA) i V_k (matricea de transformare bazată pe PCA).

Scădem media coloanelor din `train_mat` din vectorul imaginii `im`. Aceasta are scopul de a centra imaginea în jurul originii.

Transformăm baza imaginii prin înmul irea cu matricea V_k . Aceasta realizează proiec ia imaginii într-un nou spa iu bazat pe componente principale.

Calculăm eticheta prezisă (prediction) utilizând metoda KNN cu $k = 5$. Apelăm func ia `KNN` i transmitem etichetele de antrenament (`train_val`), caracteristicile transformate prin PCA (Y), imaginea de testare (`im`) i numărul de vecini ($k = 5$).

Func ia returnează eticheta prezisă (prediction) ca rezultat al func iei.