

Sveučilište u Zagrebu
Prirodoslovno-matematički fakultet
Matematički odsjek

Primjena HOSVD-a u klasifikaciji rukom pisanih znamenki

Dorotea Rajšel, Iva Sokač

16. svibnja 2018.

seminarski rad iz kolegija Matematičko modeliranje pretraživača
mentor: prof. dr. sc. Zlatko Drmač

Sadržaj

1	Uvod	3
2	Koncepti tenzora	5
2.1	HOSVD	6
2.2	Aproksimacija tenzora koristeći HOSVD	7
2.3	Ortogonalne bazne matrice	7
3	Algoritmi	9
3.1	Algoritam 1: Klasifikacija pomoću HOSVD-a	9
3.1.1	Faza treniranja	9
3.1.2	Faza testiranja	10
3.2	Algoritam 2: Kompresija i klasifikacija pomoću HOSVD	10
3.2.1	Faza treniranja	10
3.2.2	Faza testiranja	12
4	Testiranje i rezultati	14
4.1	Skup podataka	14
4.2	Pretprocesiranje	14
4.3	Rezultati testiranja prvog algoritam	14
4.4	Rezultati testiranja drugog algoritma	14
5	Zaključak	17

Sažetak

U ovom seminaru bavit ćemo se s dva algoritma za klasifikaciju rukom pisanih znamenki temeljenih na singularnoj dekompoziciji višeg reda (HOSVD). Prvi algoritam koristi HOSVD za konstrukciju modela klase i postiže rezultate klasifikacije sa pogreškom manjom od 6%. Drugi algoritam koristi HOSVD za tenzorsku aproksimaciju istodobno u dva moda. Greške klasifikacije kod drugog algoritma su skoro na 5% iako aproksimacijom reduciramo polazne podatke više od 98% prije nego li konstruiramo model klase. Stvarna klasifikacija u testnoj fazi za oba algoritma se sastoji od rješavanja niza problema najmanjih kvadrata. S obzirom na količinu izračuna za prikazani test, drugi algoritam je dvostruko učinkovitiji od prvog algoritma.

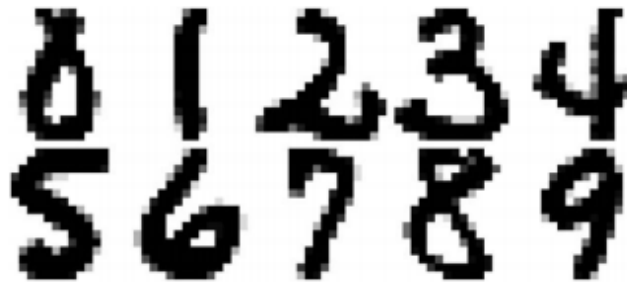
1 Uvod

Automatska klasifikacija rukom pisanih znamenaka se često smatra standardnim problemom u prepoznavanju uzoraka te uključuje mnoge poteškoće s kojima se susrećemo u ovom području. Dodjeljivanje nepoznatog objekta jednoj od 10 unaprijed zadanih klasa je težak problem za riješiti s obzirom na to da je varijacija objekata unutar iste klase velika, a istovremeno objekti iz različitih klasa mogu biti jako slični.

Postoji mnogo različitih pristupa za rješavanje ovog problema: analiza glavnih komponentata (PCA), metode potpornih vektora, metoda bliskog susjeda i k-susjedne metode, regresija, statističko modeliranje i neuronske mreže. Najbolji algoritmi se baziraju na neuronskim mrežama te tzv. tangencijalnim udaljenostima koje koriste udaljenost kao mjeru invarijantna na lokalne afine transformacije. Generalno, algoritmi s dobrim performansima imaju ili veliku opisnu složenost ili računalno veliku složenost.

U ovom radu prezentirat ćemo dva jednostavna i učinkovita algoritma s poprilično dobrim rezultatima. Oba algoritma se baziraju na singularnoj dekompoziciji višeg reda (HOSVD) tenzora. Prvi algoritam koristi HOSVD kako bi izračunao mali skup baznih matrica koje razapinju dominantan potprostor za svaku klasu znamenki. Bazne matrice se koriste kako bi se opisala nepoznata znamenka. Drugi algoritam koristi HOSVD za kompresiju skupa podataka za treniranje. Modeli klasa (ovdje bazni vektori) se računaju korištenjem samo reduciranih podataka, a sama klasifikacija se provodi kao u prvom algoritmu. Prednosti su dvostruke: opis modela klasa zahtijeva manje memorije i sama faza klasifikacije je učinkovitija bez nedostataka u samom procesu izvršavanja algoritma. Algoritam daje stopu pogreške od 5% čak i nakon 98%-tne redukcija skupa podataka za treniranje.

Posljednjih godina primjena tenzorskih metoda na probleme prepoznavanja uzoraka i u drugim područjima privlači sve više i više pažnje. Pod tenzorima podrazumijevamo višedimenzionalna ili višemodna polja. Često podaci imaju višedimenzionalnu strukturu pa je na neki način neprirodno ograničavati ih u matrice ili vektore. Kao jednostavni primjer, uzmimo vremenski niz slika. Svaka slika je dvodimenzionalno podatkovno polje i zajedno sa slikama iz različitih vremenskih koraka podaci tvore tenzor trećeg reda (trodimenzionalno podatkovno polje). U mnogim slučajevima učinkovitije je koristiti prikupljene podatke bez narušavanja njihove višedimenzionalne strukture. Tenzorska metoda se već dulje vrijeme koristi u kemijskim mjerenjima i psihometrijskim metodama. Od nedavno HOSVD nalazi primjenu i kod prepoznavanja lica.



Slika 1: Primjeri rukom pisanih znamenki iz USPS baze



Slika 2: Prikaz dijelova USPS baze znamenki

U ovom seminaru koristit ćemo rukom pisane znamenke iz američke poštanske uslužne baze podataka kako bismo testirali naša dva algoritma. Znamenke su reprezentirane slikama velicine 16×16 (sivi tonovi) nastalih skeniranjem poštanskih omotnica. Ovaj skup podataka se učestalo koristi pri procjeni klasifikacijskih algoritama. Uzorci su prikazani na slici 1.

Ostatak rada je strukturiran na sljedeći način. Drugi odjeljak sadrži sažetak rezultata o tenzorima i neke teorijske rezultate koje koristimo u algoritmima. U trećem odjeljku prezentiramo oba algoritma. Numerički testovi su opisani u četvrtom odjeljku te su detaljnije opisani dati podatkovni skupovi.

Algoritmi će biti opisani koristeći pseudo-MATLAB kod. Dakle, u primjerima kada ćemo koristiti notaciju $A(i, j, k)$ za a_{ijk} . Također, u formulama ćemo nekada koristiti MATLAB-ovsku notaciju.

2 Koncepti tenzora

Neformalno govoreći, tenzor N -tog reda je objekt s N indeksa. Dimenziju tenzora će određivati broj modova. Vektore i matrice možemo smatrati tenzorima prvog, odnosno drugog reda. Za potrebe ovog seminara bavit ćemo se slučajem kada je $N = 3$. Dakle, jednostavnosti radi, neki teorijski rezultati u ovom odjeljku bit će iskazani samo za tenzore trećeg reda $\mathcal{A} \in \mathbb{R}^{I \times J \times K}$, gdje su I, J, K pozitivni cijeli brojevi, a vektorski prostor $\mathbb{R}^{I \times J \times K}$ je dimenzije IJK . Generalizacija tenzora do proizvoljnog reda je trivijalna. Neka je vektorskom prostoru $\mathbb{R}^{I \times J \times K}$ pridružena standardna Euklidska geometrija. Definiramo skalarni produkt $\langle \mathcal{A}, \mathcal{B} \rangle$ dvaju tenzora $\mathcal{A}, \mathcal{B} \in \mathbb{R}^{I \times J \times K}$ kao

$$\langle \mathcal{A}, \mathcal{B} \rangle = \sum_{i=1}^I \sum_{j=1}^J \sum_{k=1}^K a_{ijk} b_{ijk}$$

. Kažemo da su dva tenzora \mathcal{A} i \mathcal{B} ortogonalna ako im je skalarni produkt jednak nuli, tj.

$$\langle \mathcal{A}, \mathcal{B} \rangle = 0$$

. Normu tenzora \mathcal{A} definiramo kao $\|\mathcal{A}\| = \sqrt{\langle \mathcal{A}, \mathcal{A} \rangle}$.

Analogno definiramo skalarni produkt i normu vektora i matrica. Nekada je pogodno preurediti elemente tenzora tako da dobijemo matricu. Taj proces ćemo zvati matricizacija tenzora. Transformacija tenzora \mathcal{K} u matricu u modu n je operacija gdje se nit u modu n tenzora \mathcal{K} reprezentira odgovarajućim stupcem matrice $K_{(n)}$. Stupce matrice $K_{(n)}$ možemo orijentirati ciklički prema naprijed. Tada je matricizacija tenzora trećeg reda $\mathcal{A} \in \mathbb{R}^{I \times J \times K}$ definirana određivanjem položaja u matrici elementa iz \mathcal{A} :

$$\begin{aligned} \mathbb{R}^{I \times JK} \ni A_{(1)} : a_{ijk} &= a_{iv}^{(1)}, v = j + (k - 1)K, \\ \mathbb{R}^{J \times IK} \ni A_{(2)} : a_{ijk} &= a_{jv}^{(2)}, v = k + (i - 1)I, \\ \mathbb{R}^{K \times IJ} \ni A_{(3)} : a_{ijk} &= a_{kv}^{(3)}, v = i + (j - 1)J \end{aligned}$$

Uočimo da su stupčani vektori od $A_{(n)}$ zapravo niti u modu n od \mathcal{A} .

Nema jedinstvenog načina za generalizaciju ranga za višedimenzionalne tenzore preko definicije ranga za matrice. Jedna od mogućnosti je da definiramo n -rang tenzora \mathcal{A} kao dimenziju potprostora razapetog s nitima u modu n tj. $\text{rang}_n(\mathcal{A}) = \text{rang}(A_{(n)})$, gdje je $A_{(n)}$ matrična transformacija tenzora \mathcal{A} u modu n i "rang" bez subskripta označava matični rang. Različiti n -rangovi tenzora trećeg reda (ili višeg) obično nisu jednaki kao što je to slučaj s matricama. Sada ćemo dati opću definiciju tenzorsko-matričnog množenja.

Definicija 2.1 (množenje tenzora i matrice u modu n). Za tenzor $\mathcal{A} \in \mathbb{R}^{I_1 \times \dots \times I_N}$ i matricu $F \in \mathbb{R}^{J_n \times I_n}$ definiramo produkt od \mathcal{A} i F u modu n u oznaci $\mathcal{A} \times_n F$ kao

$$(\mathcal{A} \times_n F)(i_1, \dots, i_{n-1}, j_n, i_{n+1}, \dots, i_N) = \sum_{i_n=1}^{I_n} \mathcal{A}(i_1, \dots, i_N) F(j_n, i_n)$$

Primjerice, množenje tenzora $\mathcal{A} \in \mathbb{R}^{I \times J \times K}$ u modu 1 s matricom $F \in \mathbb{R}^{L \times I}$ je dano s

$$\mathbb{R}^{L \times J \times K} \ni \mathcal{B} = \mathcal{A} \times_1 F, \mathcal{B}(l, j, k) = \sum_{i=1}^I \mathcal{A}(i, j, k) F(l, i)$$

Neka su dani tenzor $\mathcal{A} \in \mathbb{R}^{I \times J \times K}$ i matrice $F \in \mathbb{R}^{L \times I}$, $G \in \mathbb{R}^{M \times J}$ i $H \in \mathbb{R}^{N \times L}$. Tada tenzorsko-matrično množenje zadovoljava iduća svojstva:

$$\begin{aligned} (\mathcal{A} \times_1 F) \times_2 G &= (\mathcal{A} \times_2 G) \times_1 F = \mathcal{A} \times_1 F \times_2 G \in \mathbb{R}^{L \times M \times K} \\ (\mathcal{A} \times_1 F) \times_1 H &= \mathcal{A} \times_1 (HF) \in \mathbb{R}^{N \times J \times K} \end{aligned}$$

Kako su stupčani vektori od $A_{(n)}$ niti u modu n od tenzora \mathcal{A} , slijedi da množenje u modu n $\mathcal{B} = \mathcal{A} \times_n X$ može biti izraženo u terminima matriciziranog tenzora kao obično matrično množenje $B_{(n)} = X A_{(n)}$, iz čega opet složimo $B_{(n)}$ u tenzor \mathcal{B} .

2.1 HOSVD

Dekompozicija singularnih vrijednosti matrice (SVD) je vrlo koristan alat u mnogim primjenama. Bez smanjenja općenitosti, pretpostavljamo da je $F \in \mathbb{R}^{m \times n}$ za $m \geq n$.

Teorem 2.2 (Matrični SVD). *Svaka matrica $F \in \mathbb{R}^{m \times n}$ se može napisati kao produkt $F = U \Sigma V^T$, pri čemu su $U \in \mathbb{R}^{m \times m}$ i $V \in \mathbb{R}^{n \times n}$ ortogonalne matrice, a Σ je $(m \times n)$ dijagonalna matrica s nenegativnim elementima poredanim u opadajućem poretku $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n \geq 0$. Stupce matrica U i V zovemo lijevi odnosno desni singularni vektori, a σ_i su singularne vrijednosti.*

Ako promatramo matricu kao tenzor drugog reda, SVD možemo izraziti u terminima n-mod produkta $F = \Sigma \times_1 U \times_2 V$.

Slijedi generalizacija SVD-a u smislu tenzora, u nazivu HOSVD. Navodimo rezultat za tenzore reda $N=3$.

Teorem 2.3 (HOSVD). *Tenzor $\mathcal{A} \in \mathbb{R}^{I \times J \times K}$ reda 3 se može napisati kao produkt $\mathcal{A} = \mathcal{S} \times_1 U \times_2 V \times_3 W$ sa sljedećim svojstvima:*

- (1) $U \in \mathbb{R}^{I \times I}$, $V \in \mathbb{R}^{J \times J}$ i $W \in \mathbb{R}^{K \times K}$ su ortogonalne matrice
- (2) \mathcal{S} je realan tenzor iste dimenzije kao i tenzor \mathcal{A} te zadovoljava:
 - (a) Proizvoljna dva različita "slice-a" fiksirana u istom modu su ortogonalna,

$$\begin{aligned} \langle \mathcal{S}(\nu, :, :), \mathcal{S}(\lambda, :, :) \rangle &= 0 \text{ kada } \nu \neq \lambda \\ \langle \mathcal{S}(:, \nu, :), \mathcal{S}(:, \lambda, :) \rangle &= 0 \text{ kada } \nu \neq \lambda \\ \langle \mathcal{S}(:, :, \nu), \mathcal{S}(:, :, \lambda) \rangle &= 0 \text{ kada } \nu \neq \lambda \end{aligned}$$

- (b) Norme "slice-ova" u svakom modu su uređene, npr. za prvi mod imamo

$$\|\mathcal{S}(1, :, :)\| \geq \|\mathcal{S}(2, :, :)\| \geq \dots \geq 0$$

Norme u zadnjim nejednakostima su zapravo singularne vrijednosti $\sigma_i^{(n)}$ matričnog tenzora $A_{(n)}$. Neformalno govoreći, svojstvo (b) kaže da je "energija" ili "masa" jezgre tenzora \mathcal{S} koncentrirana u okolini koordinate (1,1,1). To je svojstvo koje

omogućava da koristimo HOSVD za kompresiju podataka.

Izračun HOSVD-a tenzora $\mathcal{A} \in \mathbb{R}^{I \times J \times K}$ se može obaviti tako da odvojeno računamo ortogonalne matrice U , V i W kao lijeve singularne matrice od $A_{(n)}$, $n = 1, 2, 3$:

(1) Izračunaj tri SVD-a

$$\begin{aligned} A_{(1)} &= U S^{(1)} (V^{(1)})^T \\ A_{(2)} &= V S^{(2)} (V^{(2)})^T \\ A_{(3)} &= W S^{(3)} (V^{(3)})^T \end{aligned}$$

bez eksplicitnog formiranja $V^{(j)}$ -ova

(2) Izračunaj jezgreni tenzor $\mathcal{S} = \mathcal{A} \times_1 U^T \times_2 V^T \times_3 W^T$

Uočimo da izbjegavanjem izračunavanja desnih singularnih matrica možemo iskoristiti učestalu situaciju kada je na primjer $I \ll JK$ i uštediti veliki broj operacija s pomičnim zarezom (floating point operation)

2.2 Aproksimacija tenzora koristeći HOSVD

Kod matrica kompresija podataka se često vrši aproksimacijom nižeg ranga koja se dobiva iz SVD-a, tj.

$$F = U \Sigma V^T \approx U_k \Sigma_k V_k$$

gdje je $\Sigma_k = \text{diag}(\sigma_1, \dots, \sigma_k)$, a U_k i V_k sadrže prvih k stupaca od U , odnosno V . Aproksimacija je simbolički prikazana na slici 3.

Aproksimaciju tenzora na analogan način dobijemo pomoću HOSVD-a, zahvaljujući svojstvu: norme "slice-ova" u svakom modu su uređene. Za tenzore trećeg reda aproksimacija se može napisati kao

$$\mathbb{R}^{I \times J \times K} \ni \mathcal{A} = \mathcal{S} \times_1 U \times_2 V \times_3 W \approx \tilde{\mathcal{S}} \times_1 \tilde{U} \times_2 \tilde{V} \times_3 \tilde{W},$$

gdje je $\tilde{\mathcal{S}} \in \mathbb{R}^{k_1 \times k_2 \times k_3}$, $\tilde{U} \in \mathbb{R}^{I \times k_1}$, $\tilde{V} \in \mathbb{R}^{J \times k_2}$ i $\tilde{W} \in \mathbb{R}^{K \times k_3}$. Mjera kvalitete aproksimacije može se dobiti ispitivanjem opadanja singularnih vrijednosti u modu n tenzora \mathcal{A} : greška aproksimacije je mala ako je norma izostavljenog dijela jezgre tenzora \mathcal{S} mala. Shematski prikaz aproksimacije tenzora pomoću nižeg ranga je dan na slici 3.

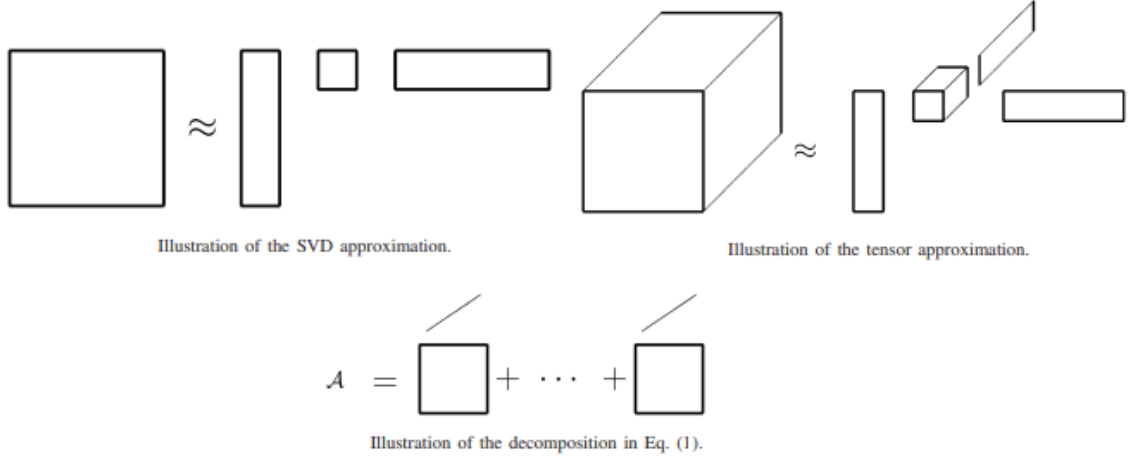
2.3 Ortogonalne bazne matrice

Matrica F se može napisati kao suma matrica ranga 1 u smislu SVD-a, tj.

$$F = U \Sigma V^T = \sum_{i=1}^n \sigma_i u_i v_i^T$$

Slična dekompozicija se može postići za tenzore trećeg reda,

$$\mathcal{A} = \sum_{\nu=1}^K A_\nu \times_3 w_\nu, \quad (1)$$



Slika 3: Ilustracija SVD dekompozicije, aproksimacije tenzora i dekompozicije (1)

gdje je

$$A_\nu = \mathcal{S}(:, :, \nu) \times_1 U \times_2 V.$$

Uočimo da je \times_3 vanjski produkt matrice A_ν i vektora w_ν pridružen tenzoru trećeg reda.

Kako je tenzor \mathcal{S} ortogonalan, slijedi da je i A_ν ortogonalna:

$$\begin{aligned} \langle A_\nu, A_\mu \rangle &= \text{tr}(A_\nu^T, A_\mu) \\ &= \text{tr}((U\mathcal{S}(:, :, \mu)V^T)^T(U\mathcal{S}(:, :, \nu)V^T)) \\ &= \text{tr}(V\mathcal{S}(:, :, \nu)^T\mathcal{S}(:, :, \mu)V^T) \\ &= \langle \mathcal{S}(:, :, \nu), \mathcal{S}(:, :, \mu) \rangle \\ &= 0, \end{aligned}$$

kada $\nu \neq \mu$ (tr označava trag matrice).

Ove ortogonalne matrice se mogu interpretirati kao skup linearno nezavisnih baznih matrica.

Ilustracija je dana na slici 3. Kvadrati reprezentiraju A_ν , a nakošene linije iznad njih vektori w_ν koji predstavljaju vanjski produkt u trećem modu.

3 Algoritmi

3.1 Algoritam 1: Klasifikacija pomoću HOSVD-a

U ovom odjeljku opisujemo kako pomoću HOSVD-a konstruiramo algoritam za klasifikaciju rukom pisanih znamenki. Skup podataka za treniranje se "ručno" klasificira. Shvaćajući svaku znamenku kao točku u $\mathbb{R}^{16 \times 16}$, prirodno je pretpostaviti da znamenke skupa za treniranje formiraju 10 dobro separiranih klastera. Inače bilo koji algoritam klasifikacije ne bi dobro radio. Istovremeno, "dominantni" vektori u svakom klasteru razapinju podskup u \mathbb{R}^{256} . Koristimo varijantu SIMCA algoritma koja konstruira mali skup ortogonalnih matrica (vektora) za svaku klasu koji razapinju dominantne dijelove odgovarajućeg potprostora. Zatim utvrđujemo koja baza najbolje opisuje nepoznatu znamenku, tj. izračunavamo aproksimacijsku grešku za svaku od 10 baza. Izračun različitih skupova baznih matrica se implementira pomoću HOSVD-a.

3.1.1 Faza treniranja

Skup ortogonalnih baznih matrica se konstruira kako je opisano u odjeljku 2.3. Baza za svaku klasu se računa preko tenzora trećeg reda sa svim znamenkama skupa za treniranje iste vrste. Neka je $\mathcal{A}^{two} \in \mathbb{R}^{16 \times 16 \times K}$ tenzor dvojki i pretpostavimo da je HOSVD izračunat. Sada iz jednadžbe $\mathcal{A} = \sum_{\nu=1}^K A_{\nu} \times_3 w_{\nu}$ slijedi

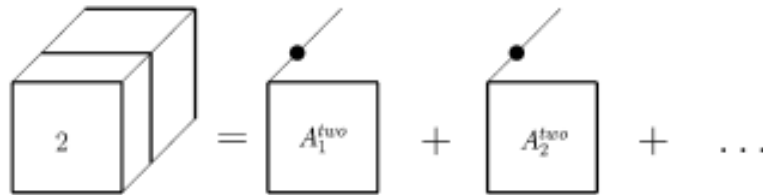
$$\mathcal{A}^{two} = \sum_{\nu=1}^K A_{\nu}^{two} \times_3 w_{\nu}^{two},$$

gdje je A_{ν}^{two} ortogonalna bazna matrica. Također, to znači da je svaka znamenka iz \mathcal{A}^{two} jedinstvena linearna kombinacija istih baznih matrica A_{ν}^{two} .

Koeficijenti u linearnoj kombinaciji su dani elementima vektora w_{ν}^{two} . Zadnju jednadžbu možemo ilustrirati kao na slici 4 gdje je proizvoljna dvojka u \mathcal{A}^{two} linearna kombinacija od A_{ν}^{two} . Koeficijenti su predstavljeni kao točke vektora u trećem modu.

Reduciramo sumu u zadnjoj jednadžbi kako bismo dobili mali i dominantni k -dimenzionalni potprostor za svaki klaster znamenki. Pretpostavimo li da su baze konstruirane za svaku klasu, od kojih svaka sadrži k baznih matrica, možemo pisati $T^{\mu} = (A_{\nu}^{\mu})_{\nu=1}^k$, $\mu = 0, 1, \dots, 9$.

Dodatno, pretpostavimo li da su bazne matrice normalizirane, slijedi $\langle A_{\nu}^{\mu}, A_{\lambda}^{\mu} \rangle = \delta_{\nu\lambda}$, gdje je $\delta_{\nu\lambda}$ Kroneckerova delta. Neka je D nepoznata znamenka i pretpostavimo da je normalizirana, tj- $\|D\| = 1$. Sad postavljamo pitanje: koji set baznih matrica T^{μ} najbolje opisuje D ?



Slika 4: Ilustracija dekompozicije tenzora dvojki

3.1.2 Faza testiranja

Promotrimo minimizacijski problem

$$\min_{\alpha_\nu^\mu} ||D - \sum_{\nu=1}^k \alpha_\nu^\mu A_\nu^\mu||,$$

gdje je μ indeks fiksne klase, a α_ν^μ su nepoznati skalari koje treba odrediti. Ovo je problem najmanjih kvadrata i lako je rješiv jer su A_ν^μ ortogonalne za fiksni μ . Rješenje je dano s

$$\bar{\alpha}_\nu^\mu = \langle D, A_\nu^\mu \rangle, \nu = 1, 2, \dots, k.$$

Zanimljivo je uočiti da je α_1^μ kosinus kuta između matrica D i A_1^μ , tj.

$$\bar{\alpha}_1^\mu = \cos(\theta^\mu) = \langle D, A_1^\mu \rangle.$$

Uvrstimo li rješenje u našu jednadžbu i iskoristimo li ortogonalnost baznih matrica, dobivamo sljedeći izraz:

$$\begin{aligned} R(\mu) &= ||D - \sum_{\nu=1}^k \bar{\alpha}_\nu^\mu A_\nu^\mu||^2 \\ &= \langle D - \sum_{\nu=1}^k \bar{\alpha}_\nu^\mu A_\nu^\mu, D - \sum_{\nu=1}^k \bar{\alpha}_\nu^\mu A_\nu^\mu \rangle \\ &= \langle D, D \rangle - \sum_{\nu=1}^k \langle D, A_\nu^\mu \rangle^2 \\ &= 1 - \sum_{\nu=1}^k \langle D, A_\nu^\mu \rangle^2. \end{aligned}$$

Sada pridružujemo D klasi za koju je $R(\mu)$ najmanji.

3.2 Algoritam 2: Kompresija i klasifikacija pomoću HOSVD

HOSVD se može iskoristiti za kompresiju podataka za treniranje prije izračuna baznih vektora za različite klase. Ovo je ideja drugog algoritma i glavni doprinos ovog rada. Također, postoji analogija s matričnom aproksimacijom nižeg ranga. Prednost u terminima računalne učinkovitosti leži u tome da znamenke iz različitih klasa projiciramo u jedan zajednički potprostor. Dakle, nepoznata znamenka se samo jednom projicira. Da imamo različite potprostore za različite klase, morali bismo projicirati nepoznatu znamenku u svaki od tih potprostora. Tada bi testna faza algoritma zahtijevala znatno veću količinu izračuna i memorije.

3.2.1 Faza treniranja

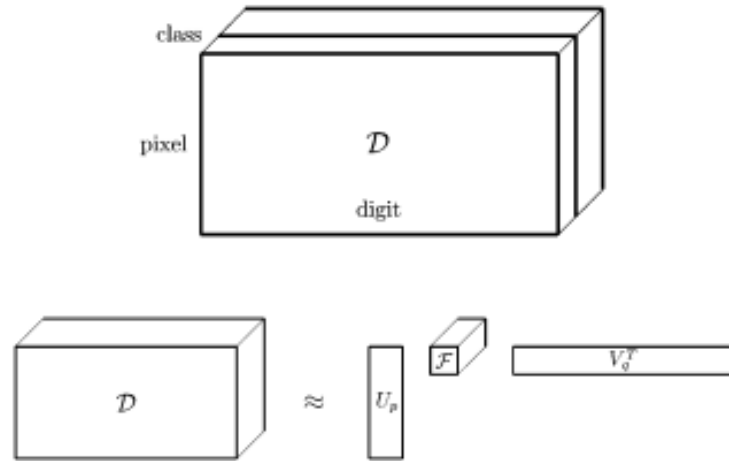
Prvo konstruiramo tenzor sa svim znamenkama iz skupa za treniranje; označimo ga s \mathcal{D} . Podaci se mogu vizualizirati kao na slici 5. Sve znamenke se pretvaraju u vektore iz \mathbb{R}^{256} . Zatim se sortiraju tako da svaki frontalni slice sadrži znamenke samo iz jedne klase. Neka \mathcal{D} ima HOSVD dan s

$$\mathcal{D} = \mathcal{S} \times_1 U \times_2 V \times_3 W \approx \mathcal{F} \times_1 U_p \times_2 V_q,$$

gdje je $U_p = U(:, 1 : p)$, $V_q = V(:, 1 : q)$ i $\mathcal{F} = \mathcal{S}(1 : p, 1 : q, :) \times_3 W$. Slika 5 prikazuje aproksimaciju tenzora.

Ovakvom aproksimacijom reduciramo prikaz individualne znamenke sa \mathbb{R}^{256} na \mathbb{R}^p i količinu znamenki u svakoj klasi na q . Reducirani tenzor \mathcal{F} se može izračunati kao

$$\mathbb{R}^{p \times q \times 10} \ni \mathcal{F} = \mathcal{D} \times_1 U_p^T \times_2 V_q^T.$$



Slika 5: Tenzor koji sadrži sve znamenke čije su dimenzije: 256 za mod piksela, 10 za mod klasa i približno 1000 za mod znamenki. Druga slika prikazuje aproksimaciju tog tenzora.

Reducirani prikaz slike možemo gledati kao stupčani prostor od U_p . Pod pretpostavkom da su p i q znatno manji od odgovarajućih dimenzija tenzora \mathcal{D} možemo značajno smanjiti originalni skup podataka za treniranje.

Tablica 2 prikazuje relativne pogreške u aproksimacijama i odgovarajuće redukcije podataka za različite p i q . Da bi aproksimacija bila dobra nužno je da izostavljeni pikseli i singularne vrijednosti znamenki budu relativno mali. Dio ovih singularnih vrijednosti je prikazan na slikama 7 i 8. Vidimo da vrijednosti piksela singularnih vrijednosti, koje odgovaraju reprezentaciji znamenki, brzo opadaju. Opadanje nije tako brzo za vrijednosti znamenki singularnih vrijednosti, što ukazuje na veću varijabilnost među različitim znamenkama iz svih klasa. Svejedno, za p i q iz tablice 2 vrijednosti pogreške su iznenađujuće male čak i u slučajevima gdje je redukcija podataka veća od 99%.

Ako aproksimaciju iz jednadžbe $\mathcal{D} = \mathcal{S} \times_1 U \times_2 V \times_3 W \approx \mathcal{F} \times_1 U_p \times_2 V_q$ napišemo kao nisko dimenzionalnu reprezentaciju tenzora znamenki u modu piksela $\mathcal{D}_p = \mathcal{D} \times_1 U_p^T = \mathcal{F} \times_2 V_q$, dobivamo iduću interpretaciju HOSVD dekompozicije. Svaki stupac vektor u \mathcal{D}_p je p element reprezentacije neke znamenke. Različiti slice-ovi od \mathcal{F} na desnoj strani se sastoje od q baznih vektora za različite klase.

Redovi od V_q su koordinate u smislu baznih vektora za različite znamenke. Zanimljivo je što jedan red od V_q daje koordinate za znamenke koje pripadaju svim 10 klasama. Prikaz je dan na slici 6.

Stupci od $\mathcal{F}^\mu := \mathcal{F}(:, :, \mu)$ predstavljaju bazne vektore za neke klase date s μ . Da bismo dobili ortogonalnost i uređenost baznih vektora za različite klase, računamo

$$\mathcal{D}_p = \mathcal{F} V_q^T$$

Slika 6

SVD od F^μ te uzimamo k najznačajnijih lijevih singularnih vektora,

$$F^\mu = [B^\mu (B^\mu)^\perp] \Sigma^\mu (Q^\mu)^T, \mu = 0, 1, \dots, 9,$$

gdje je $B^\mu \in \mathbb{R}^{p \times k}$ k -stupčana bazna matrica. Uočimo također da stupci od B^μ razapinju dominantni k -dimenzionalni potprostor od F^μ .

3.2.2 Faza testiranja

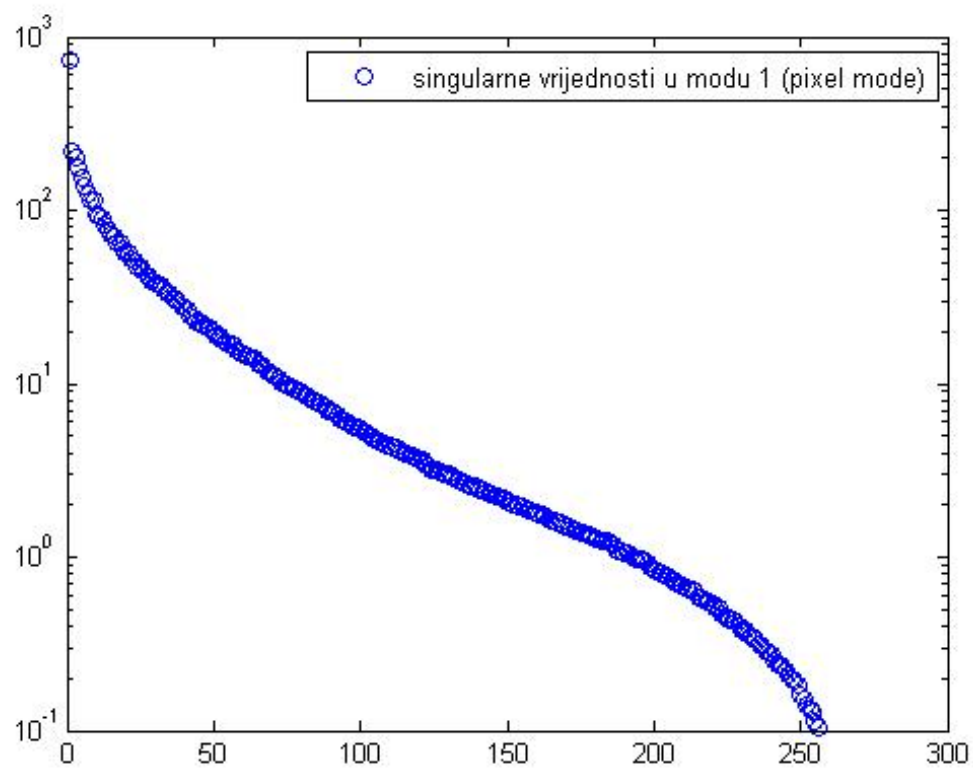
Neka je $d \in \mathbb{R}^{256}$ nepoznata znamenka. U testnoj fazi računamo nisko dimenzijsku reprezentaciju $d_p = U_p^T d$ i rješavamo idući znatno manji problem najmanjih kvadrata:

$$\min_{x^\mu} \|d_p - B^\mu x^\mu\|, \quad \text{gdje je } \mu \text{ indeks fiksne klase.}$$

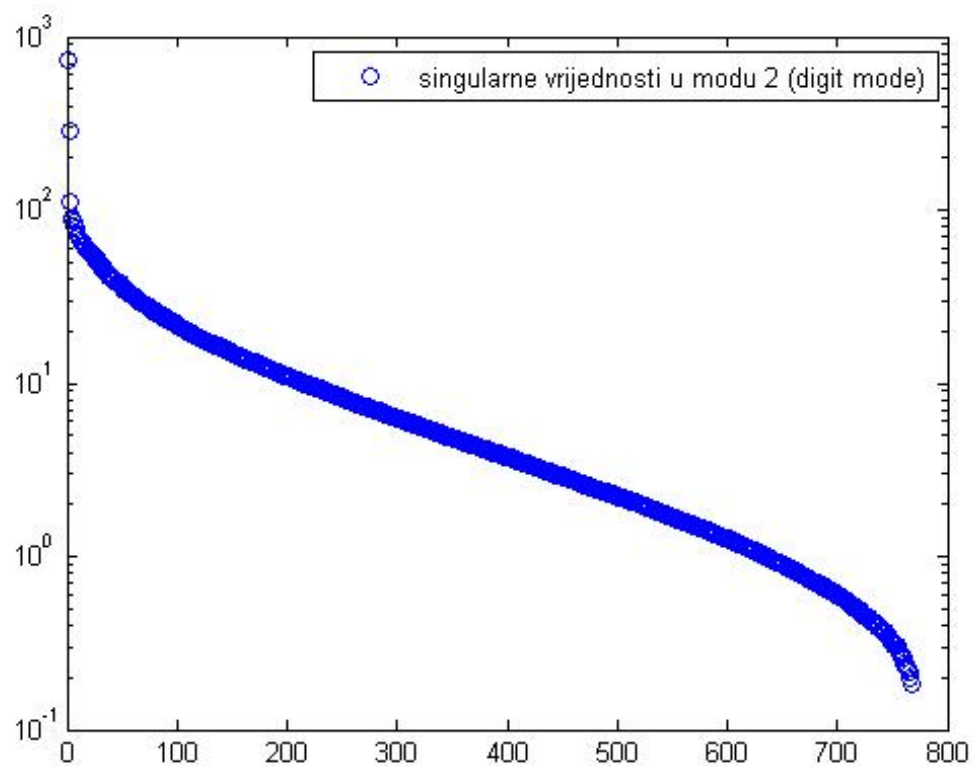
Kako su stupci od B^μ ortogonalni, rješenje je dano s $\bar{x}^\mu = (B^\mu)^T d_p$. Uvrštavanjem \bar{x}^μ u $\min_{x^\mu} \|d_p - B^\mu x^\mu\|$ dobivamo

$$R(\mu) = \|d_p - B^\mu (B^\mu)^T d_p\|.$$

Kao i ranije, indeks μ koji daje najmanji rezidual daje pretpostavljenu klasu nepoznate znamenke.



Slika 7



Slika 8

4 Testiranje i rezultati

U testiranju algoritama smo mogli varirati par parametara. Kod prvog algoritma smo varirali broj baznih matrica k , dok je kod drugog algoritma bilo moguće mijenjati broj baznih vektora k i parametre p i q koji su određivali kompresiju podataka.

4.1 Skup podataka

Skup podataka na kojem smo testirali algoritme je dostupan svima na Internetu te je često korišten za evaluaciju klasifikacijskih algoritama. Znamenke iz tog skupa su dobivene skeniranjem poštanskih brojeva s omotnica pošte US Postal. Formata su 16×16 i u sivim tonovima.

Cijeli skup sadrži 9298 znamenki. Skup je na slučajan način podijeljen na dva jednaka dijela, skup za treniranje i skup za testiranje. Ta podjela je direktno preuzeta s Interneta: www.gaussianprocess.org/gpml/data/.

Da bismo mogli koristiti cijeli skup za treniranje u algoritmima, bilo je potrebno duplicirati podatke jer znamenke nisu uniformno raspoređene u skupu.

4.2 Pretprocesiranje

Koristeći MATLAB-ove funkcije `fspecial` i `imfilter` eksperimentirale smo i sa zamučivanjem slike, u svrhu izgladivanja rubova znamenki i postizanja boljih rezultata. Koristimo Gaussian filter, zadan funkcijom

$$g(x, y) = e^{-(x^2+y^2)/(2\sigma^2)}$$

Standardna devijacija σ kontrolira zamučivanje.

4.3 Rezultati testiranja prvog algoritam

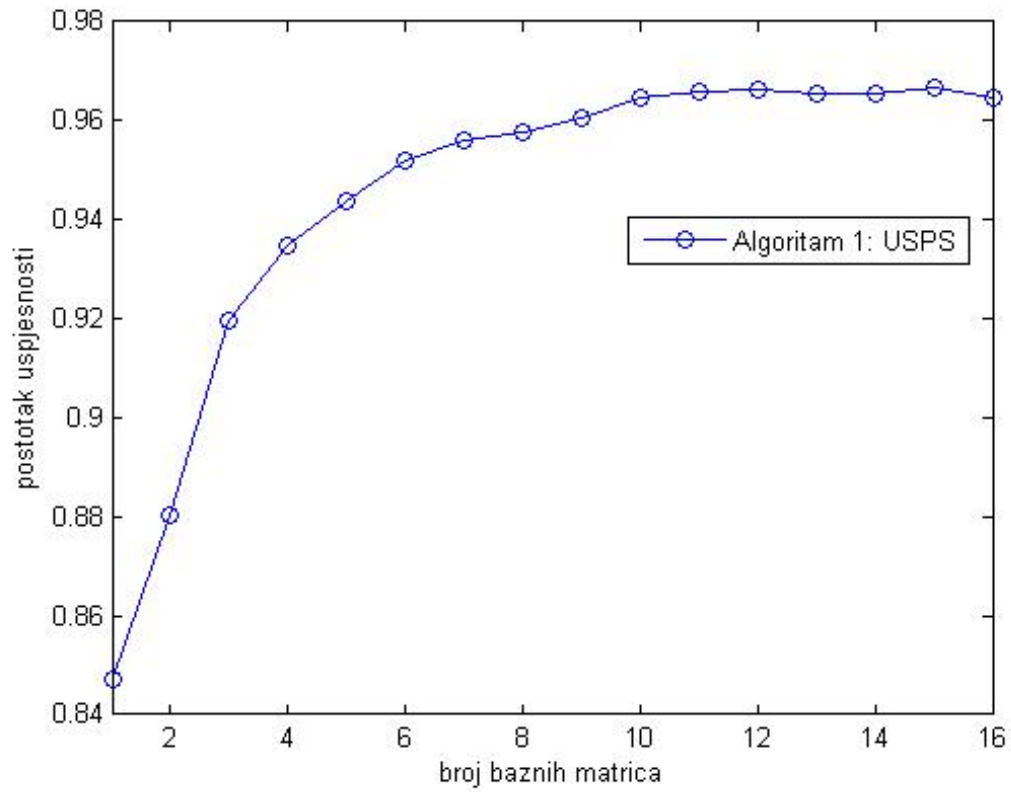
Kod testiranja prvog algoritma korišteno je najviše 16 baznih matrica, pri čemu je odabrani broj baznih matrica bio uvijek isti za sve klase znamenki. Rezultati su prikazani slikom 9.

4.4 Rezultati testiranja drugog algoritma

Varirale smo parametre p i q koji određuju redukciju podataka i broj baznih vektora k . Rezultati su prikazani tablicom 2 i slikama 10, 11 i 12. Rezultati su iznenađujuće dobri s obzirom na to da su podaci reducirani više od 98%, iako je rezultat klasifikacije neznatno lošiji od prvog algoritma.

znamenka	0	1	2	3	4	5	6	7	8	9
broj slika	767	622	475	406	409	361	420	390	377	422

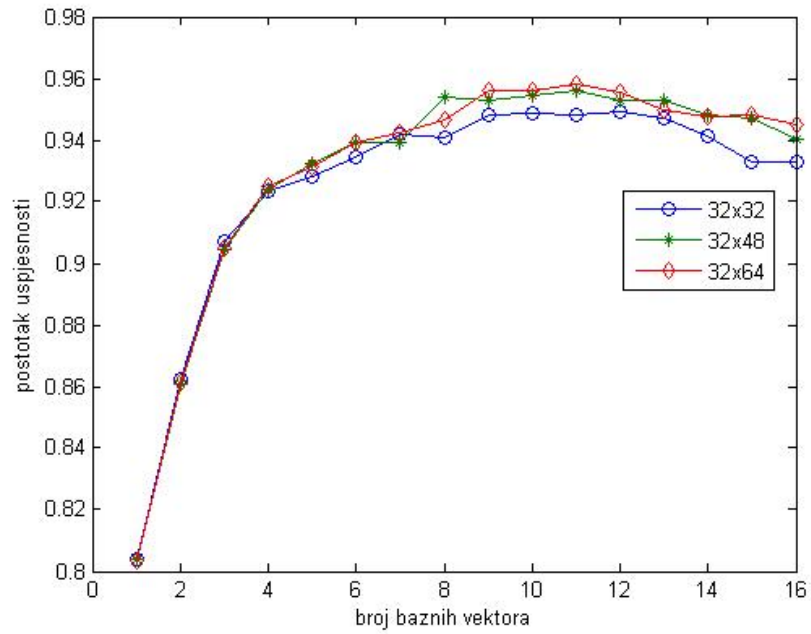
Tablica 1: Distribucija podataka skupa za treniranje



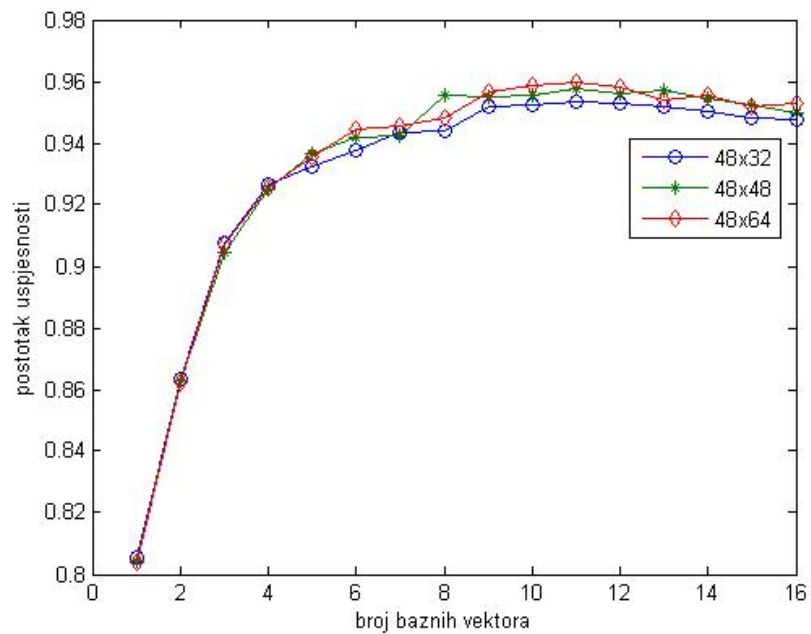
Slika 9: Rezultati klasifikacije prvim algoritmom za različite brojeve baznih matrica

p / q	32	48	64
32	94.9 (99.5)	95.6 (99.2)	95.83 (98.9)
48	95.4 (99.2)	95.7 (98.8)	95.9 (98.4)
64	95.5 (98.9)	96.1 (98.4)	96 (97.9)

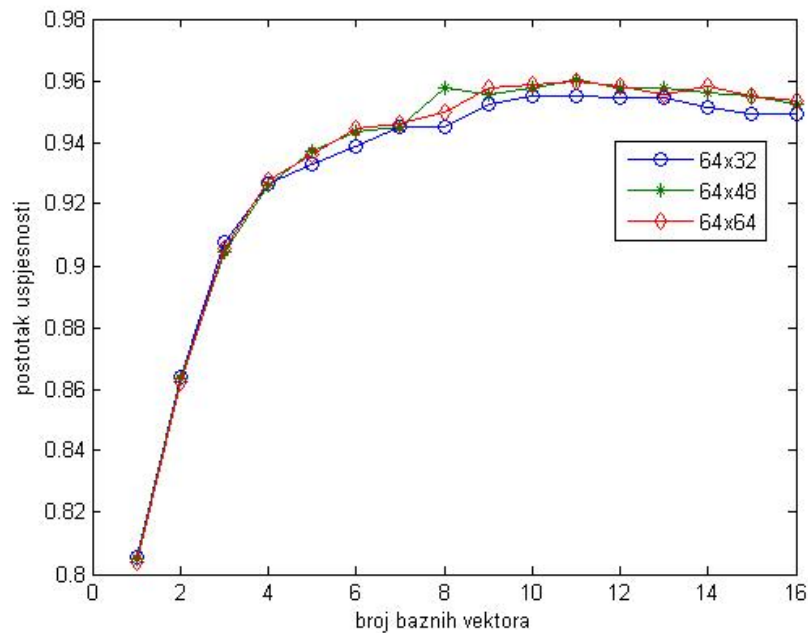
Tablica 2: Prikaz postotka točnosti (postotka redukcije podataka) za algoritam 2



Slika 10: Rezultati klasifikacije drugim algoritmom za različite brojeve baznih matrica i parametara p i q



Slika 11: Rezultati klasifikacije drugim algoritmom za različite brojeve baznih matrica i parametara p i q



Slika 12: Rezultati klasifikacije drugim algoritmom za različite brojeve baznih matrica i parametara p i q

5 Zaključak

Algoritmi prezentirani u ovom seminaru su jednostavni i efikasni, bazirani na multilinearnoj algebri. Oba algoritma imaju iznimno dobre rezultate; postotak točnosti je oko 96%. Posebno se ističe drugi algoritam koji postiže točnost visoku točnost uz čak i nakon reduciranja skupa podataka za treniranje više od 99%.

Literatura

- [SE07] Berkant Savas and Lars Elden. Handwritten digit classification using higher order singular value decomposition. *Pattern Recognition*, 40(3):993–1003, 2007.
- [USP] Gaussian Processes for Machine Learning. <http://gaussianprocess.org/gpml/data/>. [mjesto preuzimanja skupa podataka korištenog u seminaru].