

# Seminar

## Tema 6: Klasifikacija slika pomoću tenzorskih dekompozicija

Lukas Baltas, Lucija Kristić, Ivan Lipovac

23. veljače 2026.

### 1 Uvod

Tenzorske metode posljednjih su godina pronašle značajnu primjenu, osobito u područjima poput kvantne mehanike i računalne dinamike, gdje se uspješno koriste niskorangirane aproksimacije. Poseban naglasak dobiva *tensor train* (TT) dekompozicija, za koju su razvijeni i algoritmi za klasifikaciju slika.

Cilj rada je pokazati da se metode izvorno razvijene za identifikaciju upravljačkih jednažbi dinamičkih sustava — poput metode *Sparse Identification of Nonlinear Dynamics* (SINDy) i njezine tenzorske verzije *Multidimensional Approximation of Nonlinear Dynamics* (MANDy) — mogu generalizirati i primijeniti na probleme nadziranog učenja, poput klasifikacije. Pristup metode temelji se na regresiji u velikim prostorima baznih funkcija, gdje se konstrukcijom baze kao tenzorskog produkta, visokodimenzijski problemi pretvaraju u računski izvedive.

U radu se pokazuje kako se pseudoinverz problema najmanjih kvadrata može izračunati izravno u TT formatu, čime se zamjenjuje iterativni pristup direktnim računanjem rješenja. Dobiveni algoritam može se interpretirati kao kernelizirana verzija metode MANDy, pri čemu se računanje implicitno provodi u vrlo visokodimenzijskim prostorima, slično idejama kvantnog strojnog učenja. Iako su kernel metode danas često u sjeni dubokih neuronskih mreža, pokazuje se da su tenzorski kerneli konkurentni u jednostavnim zadacima klasifikacije slika.

Uz taj pristup predlaže se i metoda *Alternating Ridge Regression* (ARR), inspirirana DMRG/ALS tehnikama. Umjesto eksplicitnog računanja pseudoinverza, problem se rješava jezgreno-orijentiranom ridge regresijom u TT formatu, čime se dobivaju niskorangirane reprezentacije koeficijenata uz znatno manju računalnu i memorijsku složenost.

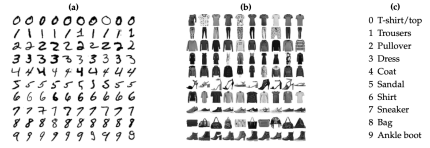


Figure 1. (a) Samples of the MNIST data set. (b) Samples of the fashion MNIST data set. Each row represents a different item type. (c) Corresponding labels for the fashion MNIST data set.

Slika 1: (a) Primjeri iz MNIST skupa podataka. (b) Primjeri iz Fashion-MNIST skupa podataka. Svaki red predstavlja različitu vrstu predmeta. (c) Pripadajuće oznake (labele) za Fashion-MNIST skup podataka.

## 2 Preduvjeti i osnovni pojmovi

### 2.1 Skupovi podataka

U radu se koriste MNIST i Fashion-MNIST skupovi podataka za zadatke klasifikacije slika. Svaka slika dimenzije  $28 \times 28$  piksela može se promatrati kao vektor

$$x \in \mathbb{R}^{784}.$$

Skupovi sadrže  $m = 60\,000$  slika za učenje i  $10\,000$  za testiranje. Oznake klasa zadane su u tzv. one-hot kodiranju te pripadaju prostoru

$$y \in \mathbb{R}^{d_0}, \quad d_0 = 10.$$

Podaci se zapisuju u obliku matrica

$$X = [x^{(1)} \quad \dots \quad x^{(m)}] \in \mathbb{R}^{d \times m}, \quad Y = [y^{(1)} \quad \dots \quad y^{(m)}] \in \mathbb{R}^{d_0 \times m}.$$

### 2.2 SINDy

Metoda SINDy izvorno je razvijena za identifikaciju dinamičkih sustava oblika

$$\dot{x} = f(x).$$

Nepoznata funkcija  $f$  aproksimira se linearnom kombinacijom unaprijed odabranih baznih funkcija.

Neka je

$$\Psi : \mathbb{R}^d \rightarrow \mathbb{R}^n$$

vektor baznih funkcija. Transformirani podaci zapisani su kao

$$\Psi_X = [\Psi(x^{(1)}) \quad \dots \quad \Psi(x^{(m)})] \in \mathbb{R}^{n \times m}.$$

Problem identifikacije svodi se na rješavanje problema najmanjih kvadrata

$$\min_{\Xi} \|Y - \Xi^\top \Psi_X\|_F,$$

gdje je  $\Xi \in \mathbb{R}^{n \times d_0}$  matrica koeficijenata. Rješenje minimalne Frobeniusove norme dano je pomoću pseudoinverza

$$\Xi^\top = Y\Psi_X^+.$$

### 2.3 MANDy

Metoda MANDy (*Multidimensional Approximation of Nonlinear Dynamics*) predstavlja tenzorsku reformulaciju metode SINDy. Umjesto eksplicitne konstrukcije matrice značajki  $\Psi_X \in \mathbb{R}^{n \times m}$ , transformacija se definira kao tenzorski produkt jednostavnih baznih funkcija.

Neka su

$$\psi_\mu : \mathbb{R}^d \rightarrow \mathbb{R}^{n_\mu}, \quad \mu = 1, \dots, p,$$

bazne funkcije. Tada se transformacija ulaza  $x \in \mathbb{R}^d$  definira kao

$$\Psi(x) = \psi_1(x) \otimes \psi_2(x) \otimes \dots \otimes \psi_p(x) \in \mathbb{R}^{n_1 \times \dots \times n_p}.$$

Dimenzija prostora značajki tada iznosi

$$n_1 \dots n_p,$$

što u općem slučaju raste eksponencijalno s brojem faktora  $p$ . Zbog toga eksplicitna pohrana i manipulacija tim objektima postaje računski neizvediva.

Ideja metode MANDy jest koristiti strukturu tenzorskog produkta kako bi se takvi visokodimenzionalni objekti mogli efikasno prikazati i koristiti u regresijskim problemima oblika

$$\min_{\Xi} \|Y - \Xi^\top \Psi_X\|_F.$$

### 2.4 Tensor Train (TT) dekompozicija

Kako bi se izbjegla eksponencijalna složenost, koristi se *tensor train* (TT) dekompozicija. Neka je

$$\mathcal{T} \in \mathbb{R}^{n_1 \times \dots \times n_p}$$

tenzor reda  $p$ . Kažemo da je  $\mathcal{T}$  u TT formatu ako se može zapisati u obliku

$$\mathcal{T}(i_1, \dots, i_p) = G^{(1)}(i_1)G^{(2)}(i_2) \dots G^{(p)}(i_p),$$

gdje su

$$G^{(\mu)} \in \mathbb{R}^{r_{\mu-1} \times n_\mu \times r_\mu}, \quad r_0 = r_p = 1,$$

tzv. TT jezgre, a  $r_\mu$  su TT-rangovi.

Ekvivalentno, element tenzora može se zapisati kao

$$\mathcal{T}(i_1, \dots, i_p) = \sum_{k_1=1}^{r_1} \dots \sum_{k_{p-1}=1}^{r_{p-1}} G_{1,i_1,k_1}^{(1)} G_{k_1,i_2,k_2}^{(2)} \dots G_{k_{p-1},i_p,1}^{(p)}.$$

Broj parametara u TT reprezentaciji iznosi

$$\sum_{\mu=1}^p r_{\mu-1} n_{\mu} r_{\mu},$$

što u tipičnim slučajevima raste linearno u  $p$ , za razliku od punog zapisa koji zahtijeva  $\prod_{\mu=1}^p n_{\mu}$  parametara.

## 2.5 Primjena tenzorskih mreža u nadziranom učenju

Tenzorski optimizacijski postupci mogu se prilagoditi problemima nadziranog učenja. Ulazni vektor  $x \in \mathbb{R}^d$  najprije se preslikava u visokodimenzionalni prostor značajki pomoću transformacije

$$\Psi : \mathbb{R}^d \rightarrow \mathbb{R}^{n_1 \times \dots \times n_p},$$

nakon čega se klasifikacija provodi pomoću odluke funkcije

$$f : \mathbb{R}^d \rightarrow \mathbb{R}^{d_0}, \quad f(x) = \Xi^{\top} \Psi(x),$$

gdje je  $\Xi$  koeficijentni tenzor zapisan u TT formatu.  $i$ -ta komponenta vektora  $f(x)$  interpretira se kao vjerojatnost da slika  $x$  pripada  $i$ -toj klasi.

Jedna od mogućih transformacija definirana je kao

$$\Psi(x) = \begin{bmatrix} \cos(\alpha x_1) \\ \sin(\alpha x_1) \end{bmatrix} \otimes \begin{bmatrix} \cos(\alpha x_2) \\ \sin(\alpha x_2) \end{bmatrix} \otimes \dots \otimes \begin{bmatrix} \cos(\alpha x_d) \\ \sin(\alpha x_d) \end{bmatrix},$$

gdje je  $\alpha$  parametar. Svaki piksel preslikava se u dvodimenzionalni vektor, inspiriran konstrukcijama iz kvantne mehanike. Takva transformacija može se interpretirati i kao ugradnja podataka u visokodimenzionalni Hilbertov prostor.

Zbog tenzorske strukture,  $\Psi(x)$  ima  $2^d$  komponenti. Za MNIST sliku ( $d = 784$ ) to znači približno  $2^{784}$  baznih funkcija, što čini eksplicitnu pohranu nemogućom. Zbog toga se koristi TT reprezentacija, pri čemu je

$$\Xi \in \mathbb{R}^{2 \times \dots \times 2 \times d_0}.$$

Za učenje koeficijentnog tenzora  $\Xi$  minimizira se funkcija troška

$$\min_{\Xi} \sum_{j=1}^m \left\| y^{(j)} - \Xi^{\top} \Psi(x^{(j)}) \right\|_2^2.$$

Algoritam iterativno optimizira TT jezgre, prilagođava TT-rangove te koristi skraćeni SVD za kontrolu računske složenosti. Smanjenje TT-rangova može se interpretirati kao oblik regularizacije.

### 3 Tenzorski algoritmi za klasifikaciju

U ovom poglavlju opisuju se dva pristupa klasifikaciji temeljena na tensor train (TT) reprezentaciji:

- kernelizirana verzija metode MANDy,
- Alternating Ridge Regression (ARR).

Oba pristupa polaze od ideje preslikavanja podataka u vrlo visokodimenzionalni prostor značajki pomoću tenzorskog produkta baznih funkcija, ali se razlikuju u načinu rješavanja regresijskog problema.

#### 3.1 Dekompozicija baze i transformirani podaci

Neka je  $x \in \mathbb{R}^d$  ulazni vektor (slika), te neka su

$$\psi_\mu : \mathbb{R}^d \rightarrow \mathbb{R}^{n_\mu}, \quad \mu = 1, \dots, p$$

bazne funkcije.

Transformacija u prostor značajki definira se kao rang-1 tenzor

$$\Psi(x) = \psi_1(x) \otimes \dots \otimes \psi_p(x) \in \mathbb{R}^{n_1 \times \dots \times n_p}.$$

Za skup podataka

$$X = [x^{(1)}, \dots, x^{(m)}] \in \mathbb{R}^{d \times m}$$

konstruira se transformirani tenzor

$$\Psi_X \in \mathbb{R}^{n_1 \times \dots \times n_p \times m},$$

čiji je  $j$ -ti presjek dan s  $\Psi(x^{(j)})$ .

Zbog tenzorske strukture vrijedi

$$(\Psi_X)_{i_1, \dots, i_p, j} = \prod_{\mu=1}^p \psi_{\mu, i_\mu}(x^{(j)}).$$

Iako je dimenzija prostora značajki

$$n = n_1 \cdot \dots \cdot n_p$$

eksponencijalno velika, TT reprezentacija omogućuje učinkovito spremanje i računanje.

### 3.2 Kernelizirana MANDy metoda

Neka su  $x^{(j)} \in \mathbb{R}^d$ ,  $j = 1, \dots, m$ , ulazni uzorci, a

$$Y = [y^{(1)} \quad \dots \quad y^{(m)}] \in \mathbb{R}^{d_0 \times m}$$

matrica oznaka (npr. one-hot kodiranje,  $d_0$  je broj klasa). Za odabrane bazne funkcije

$$\psi_\mu : \mathbb{R}^d \rightarrow \mathbb{R}^{n_\mu}, \quad \mu = 1, \dots, p,$$

definira se tenzorska preslikba značajki

$$\Psi(x) = \psi_1(x) \otimes \dots \otimes \psi_p(x) \in \mathbb{R}^N, \quad N = \prod_{\mu=1}^p n_\mu.$$

Za skup podataka uvodimo matricu značajki

$$\Psi_X = [\Psi(x^{(1)}) \quad \dots \quad \Psi(x^{(m)})] \in \mathbb{R}^{N \times m}.$$

Traži se koeficijentni tenzor  $\Xi \in \mathbb{R}^{n_1 \times \dots \times n_p \times d_0}$  (ekvivalentno, njegova matricizacija  $\widehat{\Xi} \in \mathbb{R}^{N \times d_0}$ ) takav da je

$$\widehat{\Xi}^\top \Psi_X \approx Y.$$

To vodi na problem najmanjih kvadrata

$$\min_{\widehat{\Xi}} \|Y - \widehat{\Xi}^\top \Psi_X\|_F^2.$$

Minimalno-normno rješenje dano je pseudoinverzom

$$\widehat{\Xi}^\top = Y \Psi_X^+.$$

Iz identiteta

$$\Psi_X^+ = (\Psi_X^\top \Psi_X)^+ \Psi_X^\top$$

slijedi dualni (kernel) zapis

$$\widehat{\Xi}^\top = Y (\Psi_X^\top \Psi_X)^+ \Psi_X^\top.$$

Definiranjem Gramove matrice

$$G := \Psi_X^\top \Psi_X \in \mathbb{R}^{m \times m}, \quad G_{ij} = k(x^{(i)}, x^{(j)}),$$

dobiva se

$$\widehat{\Xi}^\top = Z \Psi_X^\top, \quad Z := Y G^+ \in \mathbb{R}^{d_0 \times m}.$$

Dakle, umjesto rada u prostoru dimenzije  $N = \prod_{\mu=1}^p n_\mu$ , problem se svodi na računanje i rješavanje sustava s Gramovom matricom dimenzije  $m \times m$ .

Jezgra je inducirana preslikbom  $\Psi$ :

$$k(x, x') = \langle \Psi(x), \Psi(x') \rangle,$$

a zbog tenzorske strukture vrijedi faktORIZACIJA

$$k(x, x') = \prod_{\mu=1}^p \langle \psi_{\mu}(x), \psi_{\mu}(x') \rangle.$$

U radu se za klasifikaciju slika koristi “spin-like” preslikba

$$\Psi(x) = \begin{bmatrix} \cos(\alpha x_1) \\ \sin(\alpha x_1) \end{bmatrix} \otimes \cdots \otimes \begin{bmatrix} \cos(\alpha x_d) \\ \sin(\alpha x_d) \end{bmatrix},$$

pa se jezgra pojednostavljuje u produkt kosinusnih jezgri

$$k(x, x') = \prod_{i=1}^d \cos(\alpha(x_i - x'_i)).$$

Za novi uzorak  $x$  definiramo kernel-vektor

$$k_x = \begin{bmatrix} k(x^{(1)}, x) \\ \vdots \\ k(x^{(m)}, x) \end{bmatrix} \in \mathbb{R}^m,$$

te je klasifikacijska funkcija

$$f(x) = Zk_x \in \mathbb{R}^{d_0}, \quad \hat{c}(x) = \arg \max_{i=1, \dots, d_0} f_i(x).$$

### Algoritam

1. Izračunati Gramovu matricu  $G$ , pri čemu se iskorištava produktna struktura jezgre (u radu i Hadamardovi produkti lokalnih matrica).
2. Riješiti sustav  $ZG = Y$  (ili odgovarajući least-squares problem ako je  $G$  singularna).
3. Za svaki testni uzorak izračunati  $k_x$ , zatim  $f(x) = Zk_x$ , te klasu odrediti pravilom  $\arg \max$ .

Ova metoda daje zatvoreni dualni oblik rješenja (bez sweeping iteracija po TT jezgrama), ali zahtijeva pohranu i obradu Gramove matrice veličine  $m \times m$ , pa memorijska i računalna složenost snažno rastu s brojem uzoraka.

### 3.3 Alternating Ridge Regression (ARR)

Za razliku od kernelizirane MANDy metode, ARR ne prelazi u dualni oblik, nego izravno aproksimira koeficijentni tenzor  $\Xi$  u TT formatu niskog ranga. U skladu s radom, problem se rješava *po klasama*: za svaki red matrice oznaka  $Y_{i,:}$ , traži se tenzor  $\Xi_i$  koji minimizira

$$\min_{\Xi_i} \|Y_{i,:} - \Xi_i^\top \Psi_X\|_2^2,$$

pri čemu je  $\Xi_i \in \mathbb{R}^{n_1 \times \dots \times n_p}$  zapisan u TT formatu

$$\Xi_i \equiv (G_i^{(1)}, \dots, G_i^{(p)}), \quad G_i^{(\mu)} \in \mathbb{R}^{r_{\mu-1} \times n_\mu \times r_\mu}, \quad r_0 = r_p = 1.$$

Ključna ideja ARR-a je da je ciljna funkcija nelinearna u svim TT jezgrama zajedno, ali *linearna u jednoj jezgri* ako su sve ostale fiksirane. Zato se primjenjuje ALS/DMRG-inspirirani sweeping postupak: tijekom optimizacije ažurira se jedna jezgra  $G_i^{(\mu)}$ , dok su ostale privremeno fiksirane.

Za konstrukciju lokalnog podproblema u radu se koristi modificirana TT-reprezentacija transformiranog podatkovnog tenzora (“mutable” prikaz s premještanjem dodatnog moda), što omogućuje efikasnu konstrukciju mikromatrice i smanjuje memorijsku potrošnju stackova. Podproblem za  $\mu$ -tu jezgru gradi se iz:

- lijevog stacka  $P_\mu$  (kontrakcije fiksiranih lijevih jezgri),
- odgovarajuće jezgre transformiranog podatkovnog tenzora  $\tilde{\Psi}_{X,\mu}^{(\mu)}$ ,
- desnog stacka  $Q_\mu$  (kontrakcije fiksiranih desnih jezgri).

Njihovom kontrakcijom i matricizacijom dobiva se mikromatrica  $M_\mu$ , pa se lokalno rješava mali regresijski problem (u radu u obliku  $\min_v \|w - vM_\mu\|_2$ , uz  $w = Y_{i,:}$ ).

**Algoritam** Za svaku klasu  $i = 1, \dots, d_0$  (neovisno, pa i paralelizabilno):

1. Postavi  $w = Y_{i,:}$  i inicijalni TT tenzor  $\Xi_i$ ; zatim desno-ortonormaliziraj TT jezgre.
2. Izračunaj desni stack  $Q_p, \dots, Q_1$ .
3. **Prvi sweep** (s lijeva na desno): za  $\mu = 1, \dots, p-1$ 
  - ažuriraj lijevi stack  $P_\mu$ ,
  - konstruiraj mikromatricu  $M_\mu$  iz  $P_\mu$ ,  $\tilde{\Psi}_{X,\mu}^{(\mu)}$  i  $Q_\mu$ ,
  - riješi lokalni problem i iz dobivenog rješenja izdvoji ažuriranu jezgru (QR/SVD korak).
4. **Drugi sweep** (s desna na lijevo): analogno za  $\mu = p, \dots, 1$ , uz ažuriranje desnog stacka  $Q_\mu$ .
5. Ponovi sweep-ove nekoliko puta (u radu tipično više prolaza) do zadovoljavajuće točnosti.

Trunkirani SVD pri rješavanju lokalnih podproblema istodobno stabilizira račun i ograničava TT rangove, pa djeluje kao oblik regularizacije. Nakon što se aproksimiraju tenzori  $\Xi_i$  za sve klase, ukupni koeficijentni tenzor zapisuje se kao

$$\Xi = \sum_{i=1}^{d_0} \Xi_i \otimes e_i,$$



a klasifikacijska funkcija ostaje

$$f(x) = \Xi^\top \Psi(x), \quad \hat{c}(x) = \arg \max_{i=1, \dots, d_0} f_i(x).$$

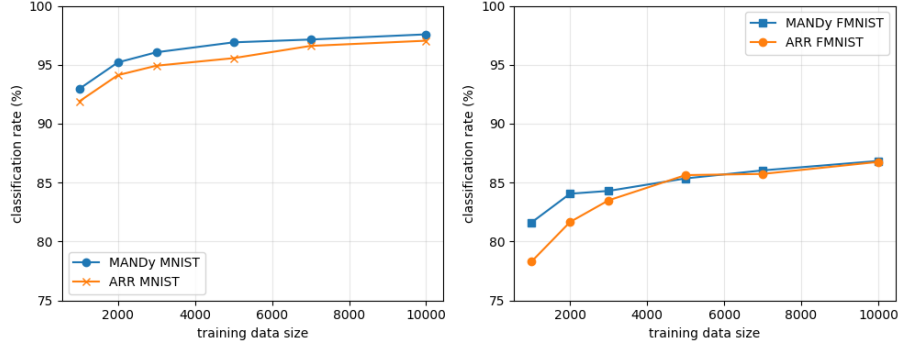
## 4 Rezultati i numerički eksperimenti

U ovom poglavlju prikazujemo eksperimentalne rezultate primjene kernelizirane MANDy metode i Alternating Ridge Regression (ARR) na skupovima MNIST i Fashion-MNIST.

Svi eksperimenti provedeni su nad podskupovima trening skupa veličine od 1000 do 10000 uzoraka, dok je testni skup fiksiran na 5000 uzoraka. Ulazne slike dimenzije  $28 \times 28$  preslikane su u vektor  $x \in \mathbb{R}^{784}$ , a zatim je primijenjena i redukcija dimenzije na 196 značajki pomoću  $2 \times 2$  downscalinga (prosjeak po blokovima piksela).

### 4.1 Learning curves

Na slici 2 prikazana je ovisnost točnosti klasifikacije o veličini trening skupa za MNIST i Fashion-MNIST.



Slika 2: Točnost klasifikacije u ovisnosti o veličini trening skupa za MNIST i Fashion-MNIST. Lijevo: MNIST. Desno: Fashion-MNIST.

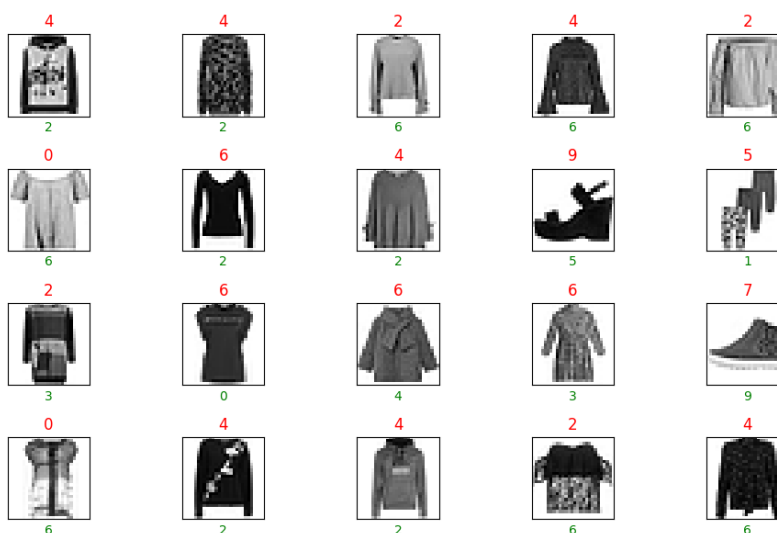
Za MNIST se vidi stabilan rast točnosti s porastom broja trening uzoraka, pri čemu kernelizirana MANDy metoda u pravilu postiže nešto višu točnost od ARR-a. To je u skladu s teorijom, jer kernelizirana varijanta koristi dualno rješenje problema najmanjih kvadrata u implicitno visokodimenzionalnom prostoru značajki, bez dodatne aproksimacije TT-rangom.

Kod Fashion-MNIST-a točnost je osjetno niža nego na MNIST-u, a razlika između metoda je manja. Razlog je veća vizualna sličnost pojedinih klasa (posebno gornjih odjevnih predmeta), pa produktna jezgra po pikselima teže razdvaja klase koje dijele slične obrise i teksture.

## 4.2 Primjeri pogrešnih klasifikacija i konfuzijske matrice

U nastavku su prikazana po dva para rezultata za MNIST i Fashion-MNIST. U svakom paru prvo je prikazana slika pogrešno klasificiranih uzoraka, a zatim pripadna konfuzijska matrica. Važno je napomenuti da su prikazane konfuzijske matrice računane nad *pogrešno klasificiranim uzorcima*, pa prikazuju strukturu pogrešaka (tj. koje se klase međusobno najčešće miješaju), a ne ukupnu konfuziju svih uzoraka.

### MNIST – primjer 1

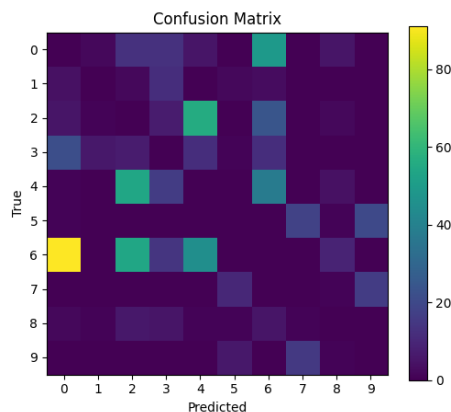


Slika 3: Primjeri pogrešno klasificiranih znamenki (MNIST), primjer 1. Crveno: predikcija modela. Zeleno: stvarna oznaka.

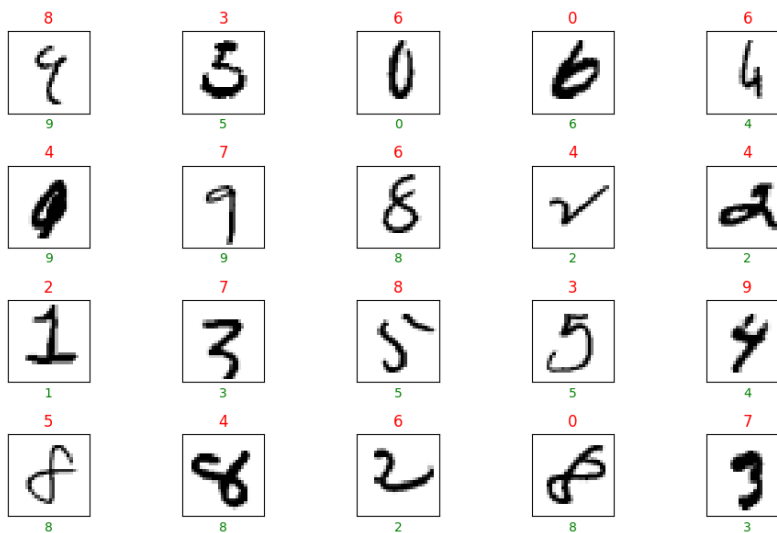
Na slici 3 vidljivo je da model najčešće griješi na vizualno dvosmislenim znamenkama, osobito kada su potezi tanki, nepotpuni ili neuobičajeno napisani. Uočljive su zabune tipa  $4 \leftrightarrow 9$ ,  $3 \leftrightarrow 5$ ,  $2 \leftrightarrow 7$  te  $0 \leftrightarrow 6/8/9$ , što je u skladu s oblicima znamenki. Konfuzijska matrica na slici 4 potvrđuje da su pogreške koncentrirane upravo između geometrijski sličnih klasa.

### MNIST – primjer 2

Drugi primjer pokazuje sličan obrazac: većina pogrešaka pojavljuje se na granicama između klasa čiji se rukopisni oblici preklapaju. Posebno su osjetljive znamenke koje ovise o lokalnim detaljima (npr. zatvorenost petlje, nagib poteza, duljina repa), dok produktivna jezgra dominantno modelira globalnu sličnost slike. Zbog toga model ne griješi nasumično, nego pretežno na uzorcima koji su i ljudskom promatraču teži za jednoznačno razdvajanje.



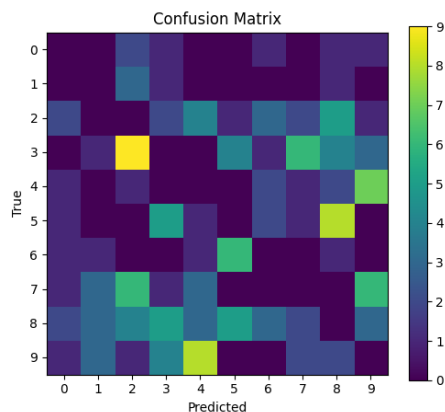
Slika 4: Konfuzijska matrica pogrešno klasificiranih uzoraka (MNIST), primjer 1.



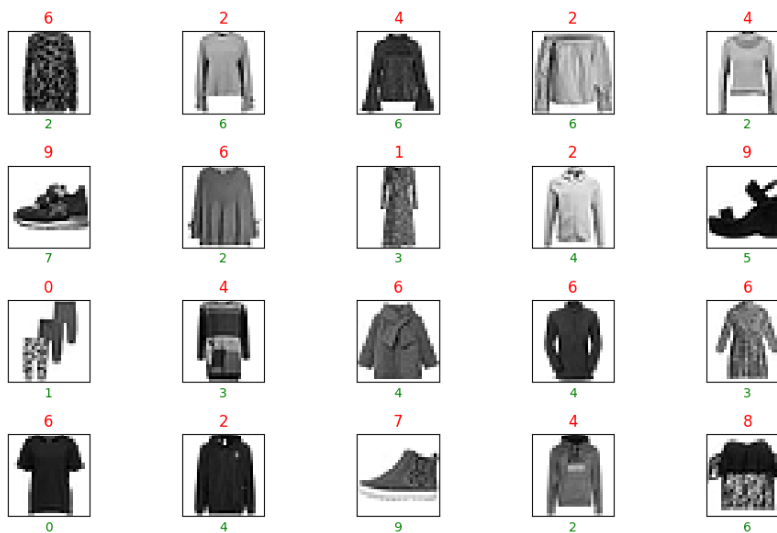
Slika 5: Primjeri pogrešno klasificiranih znamenki (MNIST), primjer 2. Crveno: predikcija modela. Zeleno: stvarna oznaka.

### Fashion-MNIST – primjer 1

Kod Fashion-MNIST-a pogreške su češće i semantički smislenije nego kod MNIST-a. Na slici 7 vidi se da model često miješa kategorije sličnog kroja i siluete, osobito među gornjim odjevnim predmetima (npr. majica, pullover, kaput, košulja), ali i među pojedinim vrstama obuće. Konfuzijska matrica na slici 8 pokazuje

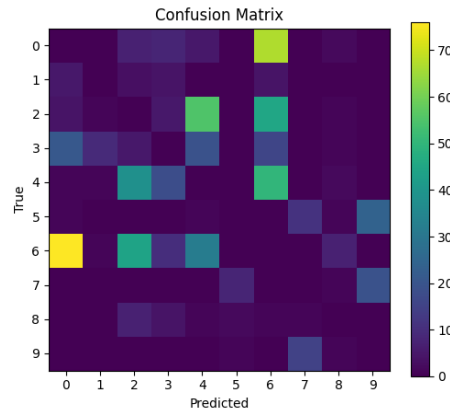


Slika 6: Konfuzijska matrica pogrešno klasificiranih uzoraka (MNIST), primjer 2.

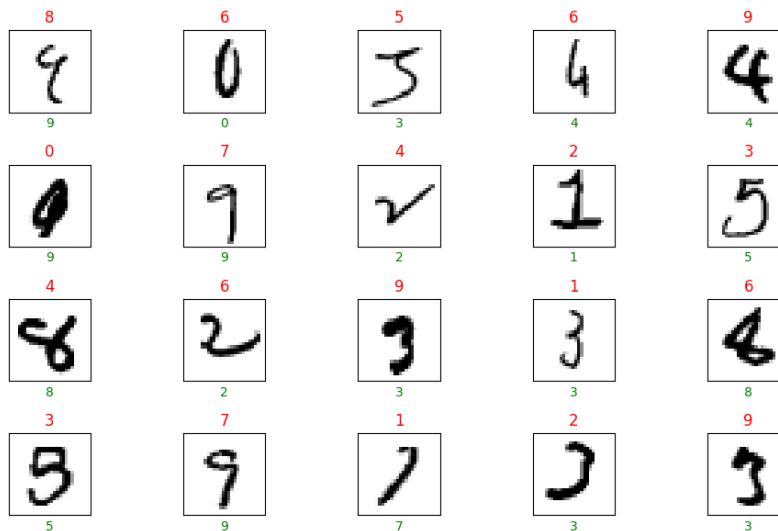


Slika 7: Primjeri pogrešno klasificiranih slika iz Fashion-MNIST skupa, primjer 1. Crveno: predikcija modela. Zeleno: stvarna oznaka.

da su pogreške koncentrirane u nekoliko “skupina” sličnih klasa, što upućuje na ograničenja modela koji radi nad globalnim pikselnim preslikavanjem bez hijerarhijske ekstrakcije značajki.



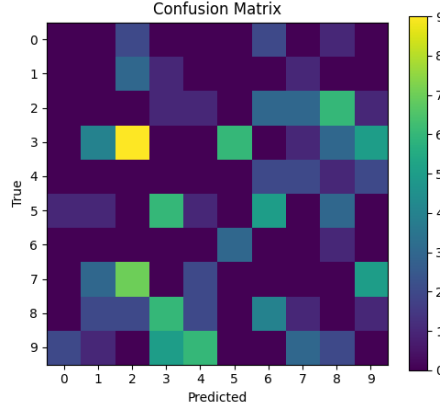
Slika 8: Konfuzijska matrica pogrešno klasificiranih uzoraka (Fashion-MNIST), primjer 1.



Slika 9: Primjeri pogrešno klasificiranih slika iz Fashion-MNIST skupa, primjer 2. Crveno: predikcija modela. Zeleno: stvarna oznaka.

## Fashion-MNIST – primjer 2

Drugi primjer dodatno potvrđuje isti trend: najproblematičnije su klase koje dijele sličan vanjski obris, dok se razlikuju u finijim detaljima teksture, rukava ili ovratnika. Slično vrijedi i za obuču, gdje se razlike često svode na lokalne detalje i profil potplata. To objašnjava zašto su rezultati na Fashion-MNIST-u



Slika 10: Konfuzijska matrica pogrešno klasificiranih uzoraka (Fashion-MNIST), primjer 2.

sustavno slabiji od rezultata na MNIST-u, iako model i dalje hvata relevantnu globalnu strukturu slike.

## 5 Zaključak

U radu smo analizirali primjenu tenzorskih metoda na klasifikaciju slika, pri čemu se podaci preslikavaju u iznimno visokodimenzionalan prostor pomoću tenzorskog produkta, a računska izvedivost postiže se TT (Tensor Train) reprezentacijom.

Rezultati pokazuju da kernelizirana MANDy metoda na skupu MNIST postiže vrlo visoku točnost i u pravilu nadmašuje ARR, što je u skladu s teorijom jer koristi dualno rješenje problema najmanjih kvadrata preko Gramove matrice. Istodobno, grafovi potvrđuju da njezina memorijska i računalna složenost raste s brojem uzoraka zbog rada s matricom dimenzije  $m \times m$ .

ARR metoda, s druge strane, izravno konstruira niskorangirani TT tenzor koeficijena i time smanjuje memorijske zahtjeve, ali zahtijeva iterativnu optimizaciju (sweep-ove) te je osjetljiva na izbor TT rangova, regularizacije i broja prolaza. To se vidi i u primjerima, gdje ARR daje nešto nižu točnost, ali zadržava dobru kvalitetu klasifikacije uz znatno kompaktniju reprezentaciju modela.

Analiza konfuzijskih matrica i pogrešno klasificiranih primjera dodatno potvrđuje da modeli najčešće griješe na vizualno sličnim klasama: kod MNIST-a na geometrijski srodnim znamenkama, a kod Fashion-MNIST-a na semantički sličnim kategorijama odjeće i obuće. Time se pokazuje da tenzorsko preslikavanje uspješno hvata globalnu strukturu uzoraka, ali teže razdvaja klase koje se razlikuju u finijim lokalnim detaljima.

Kao zaključak, tenzorske metode predstavljaju dobar pristup klasifikaciji u

visokodimenzionalnim prostorima značajki, posebno onda kada se želi izbjeći eksplicitna konstrukcija i pohrana punog prostora.