

Klasifikacija slika pomoću tenzorskih dekompozicija

Kernelizirana MANDy i Alternating Ridge Regression (ARR)

Lukas Baltas · Lucija Kristić · Ivan Lipovac

23. veljače 2026. · PMF Matematički odsjek, Zagreb

- Motivacija: klasifikacija slika kao regresija u velikom prostoru značajki
- Tenzorsko preslikavanje: $\Psi(x)$ kao tenzorski produkt jednostavnih baznih funkcija
- TT (Tensor Train) reprezentacija za računski izvediv rad
- Dva pristupa: kernelizirana MANDy i ARR (sweep optimizacija)
- Eksperimenti: MNIST i Fashion-MNIST (1k–10k train, 5k test)

Ideja: preslikavanje u veliki prostor značajki

Tenzorsko preslikavanje

Slika (28×28) se tretira kao vektor $x \in \mathbb{R}^{784}$.

Preslikavanje $\Psi(x)$ gradi se kao tenzorski produkt lokalnih funkcija, pa dimenzija prostora značajki raste eksponencijalno.

TT reprezentacija omogućuje da ne pohranjujemo svih 2^d komponenti.

Spin-like preslikavanje (po pikselima):

$$\Psi(x) = \bigotimes_{i=1}^{\{d\}} [\cos(\alpha x_i), \sin(\alpha x_i)]^T \Rightarrow k(x, x') = \prod_{i=1}^{\{d\}} \cos(\alpha (x_i - x'_i))$$

Što dobivamo?

Linearni model u prostoru $\Psi(x)$, ali bez eksplicitne konstrukcije tog prostora: ili preko kernela (MANDy) ili preko TT-učena koeficijenta (ARR).

Metode: kernel MANDy vs. ARR

Kernelizirana MANDy

- Radi u dualnom obliku preko Gramove matrice G ($m \times m$).
- Jezgra se računa produktno:
$$k(x, x') = \prod \cos(\alpha(x_i - x'_i)).$$
- Učenje: riješi $ZG = Y$
- Predikcija: $f(x) = Z k_x$; $klasa = \operatorname{argmax} f_{i(x)}$.

+ Prednost: vrlo dobra točnost za umjeren m .
– Mana: memorija i vrijeme rastu $\sim m^2$ (G) i rješavanje sustava.

Alternating Ridge Regression (ARR)

- Uči koeficijentni tenzor Ξ direktno u TT formatu niskog ranga.
- Po klasama: $\min ||Y_{\{i,: \}} - \Xi_i^T \Psi_X||^2$.
- Iterativno (sweep): fiksiraj sve jezgre osim jedne, riješi mali ridge problem, zatim QR/SVD i trunciranje ranga.

+ Prednost: manja memorija (bez pune G).
– Mana: iterativno, osjetljivo na rang/regularizaciju/broj sweepova.

Algoritam: kernelizirana MANDy

Ulaz

Trening uzorci $\{x^j\}_{j=1..m}$, oznake $Y \in R^{d_0 \times m}$, parametar α (i opcionalno λ).

Koraci

- Izračunaj Gramovu matricu G , gdje je $G_{\{ij\}} = k(x^i, x^j)$
- Riješi linearni sustav: $ZG=Y$
- Za novi x izračunaj k_x i $f(x) = Z k_x$; predikcija je $\operatorname{argmax}_i f_i(x)$

Složenost: pohrana G zahtijeva $O(m^2)$, rješavanje sustava postaje skupo za velike m .

Algoritam: Alternating Ridge Regression (ARR)

Ideja

TT jezgre se ažuriraju jedna po jedna. Lokalni problem je ridge-regresija nad mikromatricom, a TT rang se kontrolira QR/SVD koracima i trunciranjem.

Sweep postupak (po klasi i)

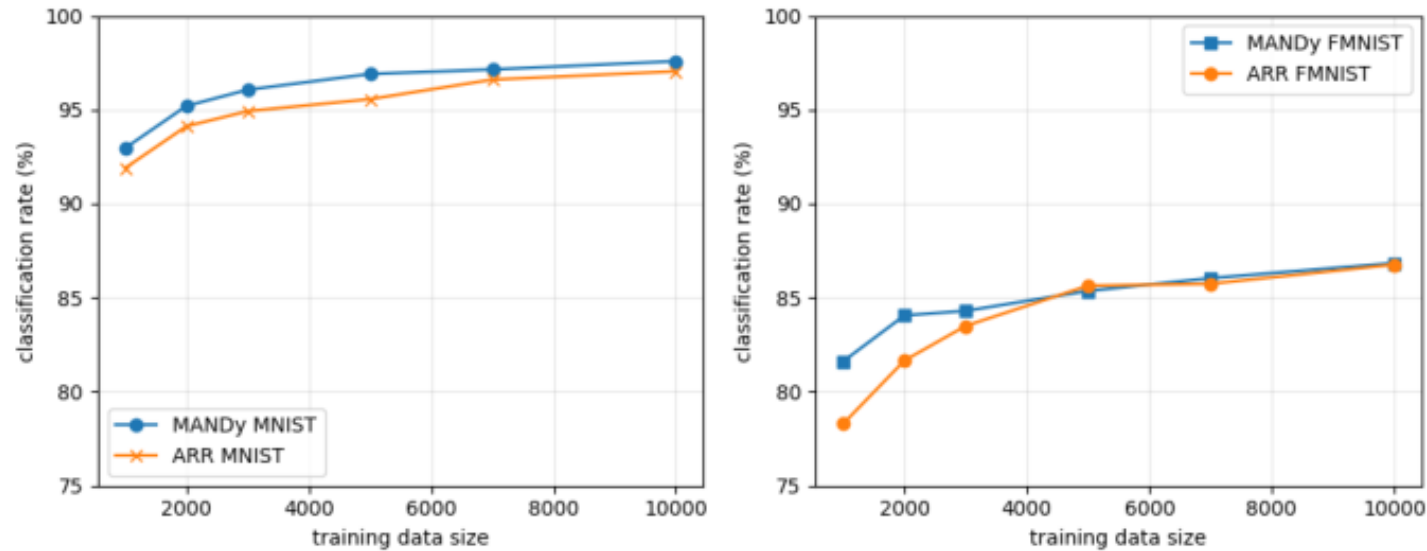
- Inicijaliziraj TT Ξ_i i desno-ortonormaliziraj jezgre.
- Izračunaj desni stack Q (kontrakcije desnih jezgri).
- Sweep $L \rightarrow R$: za svaku jezgru μ konstruiraj mikromatricu i riješi lokalni ridge problem; zatim QR/SVD i trunciranje.
- Sweep $R \rightarrow L$: analogno, uz ažuriranje desnog stacka.
- Ponovi nekoliko sweepova do stabilizacije točnosti / funkcionala.

Prednost: memorijski značajno lakše od pune Gramove matrice, ako su TT rangovi mali.

Postav

- Skupovi podataka: MNIST i Fashion-MNIST (10 klasa).
- Trening podskupovi: 1 000 \rightarrow 10 000 uzoraka (više točaka).
- Testni skup: 5 000 uzoraka (fiksno).
- Ulaz: $28 \times 28 \rightarrow$ vektor $x \in R^{784}$; downscaling $2 \times 2 \rightarrow$ 196 značajki.
- Metrike: accuracy, learning curves, primjeri pogrešaka i konfuzija pogrešno klasificiranih uzoraka.

Rezultati: learning curves

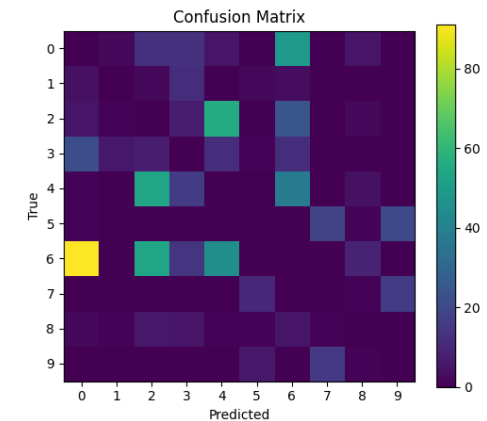
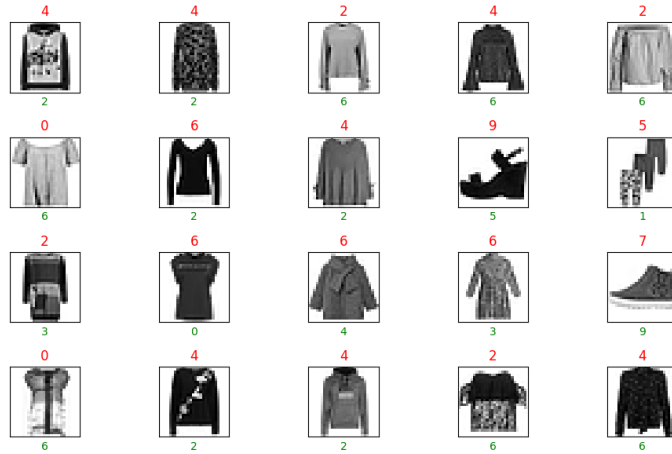


Sažetak

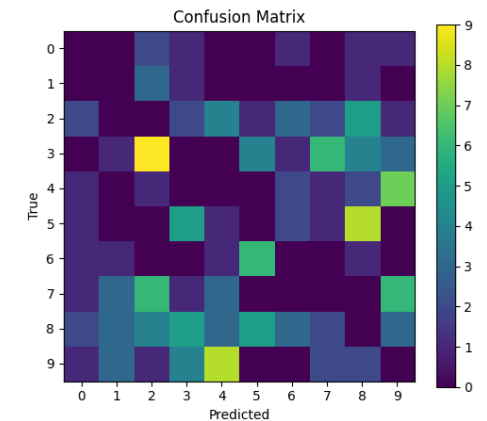
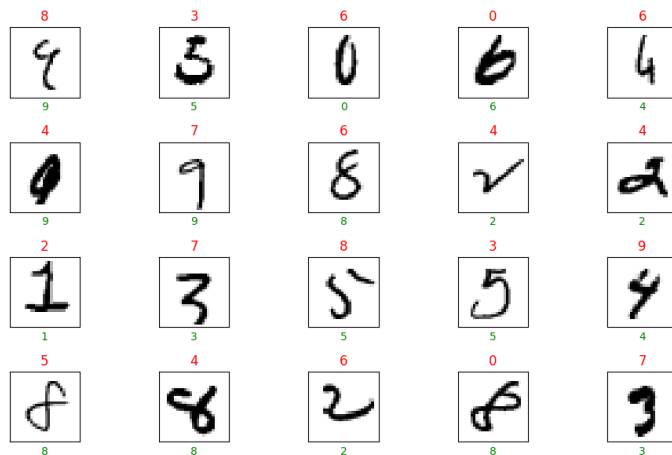
- MNIST: kernel MANDy stabilno malo bolji od ARR; ~97–98% za 10k uzoraka.
- Fashion-MNIST: niža točnost (~86–87%); metode su bliže jedna drugoj.
- Razlike su u skladu s teorijom: MANDy nema TT aproksimaciju, ARR dobiva kompaktniji model.

MNIST: pogreške i konfuzija (2 primjera)

Primjer 1



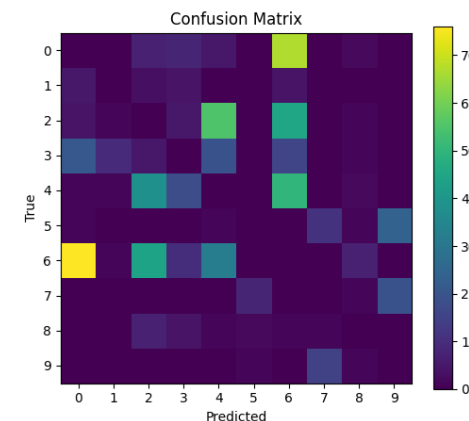
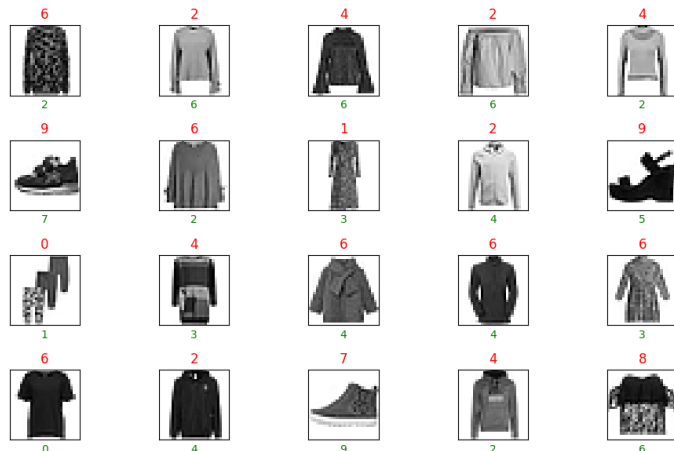
Primjer 2



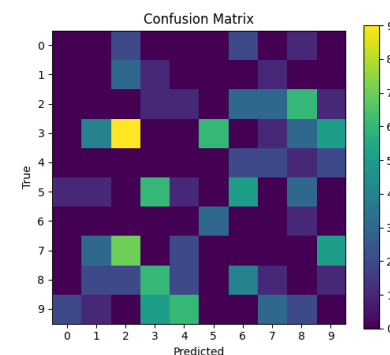
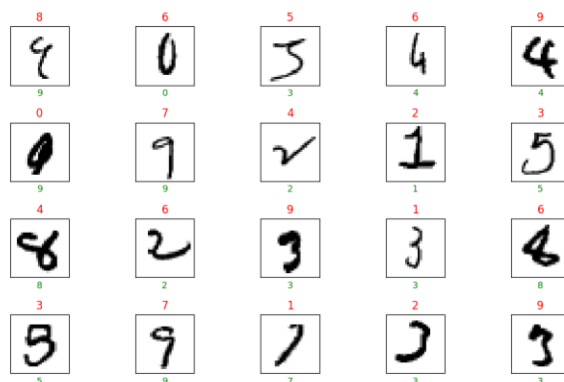
Tipične zabune: $4 \leftrightarrow 9$, $3 \leftrightarrow 5$, $2 \leftrightarrow 7$, $0 \leftrightarrow 6/8/9$ — pogreške su uglavnom između geometrijski sličnih znamenki.

Fashion-MNIST + zaključak

Primjer 1



Primjer 2



Zaključak: MANDy daje veću točnost za umjeren m, ali lošije skalira (G je $m \times m$). ARR bolje skalira preko TT niskog ranga, uz cijenu iteracija i podešavanja hiperparametara. Budući rad: bogatija preslikavanja, stabilniji ARR, hibrid s CNN.