

Primerjava pristopov učenja pri matrični faktorizaciji v priporočilnih sistemih

Marjan Meglen

30.5.2025

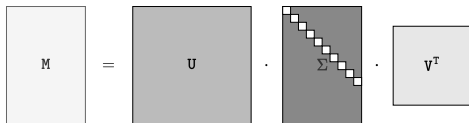
Vrste priporočilnih sistemov

- ▶ Priporočilni sistemi so del platform kot so Amazon, Netflix, Spotify.
- ▶ Njihov cilj je priporočanje vsebin na podlagi uporabniških preferenc.
- ▶ Poznamo tri glavne tipe:
 - ▶ **Vsebinski sistemi:** priporočajo glede na uporabnikove preference.
 - ▶ **Sodelovalno filtriranje:** temelji na podobnosti med uporabniki.
 - ▶ **Hibridni pristopi:** združujejo vsebinski in sodelovalni pristop.

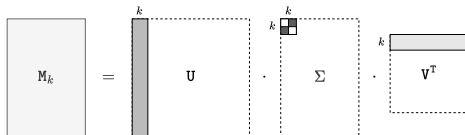
Matrična faktorizacija

- ▶ Ena izmed najpogostejše uporabljenih tehnik znotraj **sodelovalnega filtriranja**.
- ▶ Postala znana po Netflixovem tekmovanju za izboljšanje priporočilnega sistema.
- ▶ Populariziral jo je Simon Funk.
 - ▶ Algoritem **odrezani razcep na singularne vrednosti**.

Odrezani razcep na singularne vrednosti

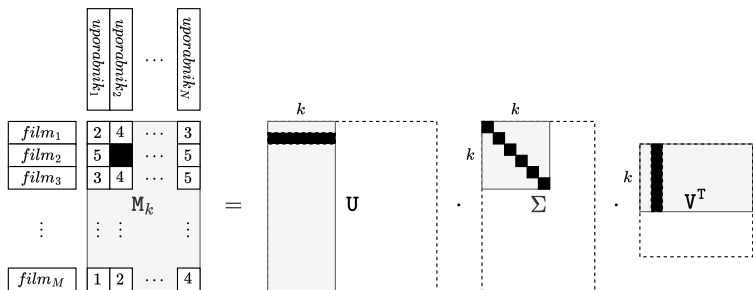


Slika: Razcep na singularne vrednosti (SVD). (Vir: <https://timbaumann.info/svd-image-compression/>)



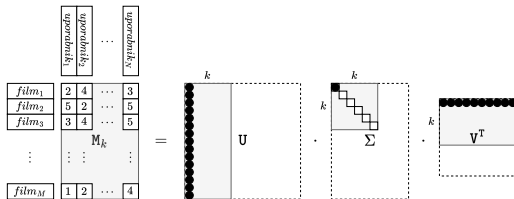
Slika: Odrezani razcep na singularne vrednosti (truncated SVD).

Napoved ocene z dobljenimi faktorji

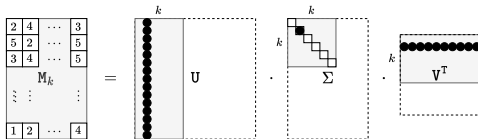


Slika: Odrezani singularni razcep matrike na primeru matrike ocen filmov.

FunkSVD

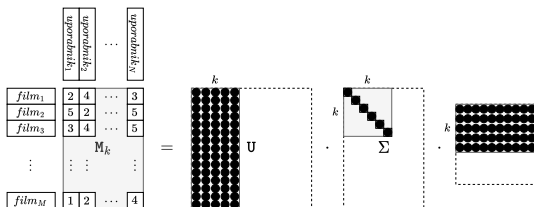


Slika: Učenje prvega latentnega faktorja.



Slika: Učenje drugega latentnega faktorja.

Sočasno učenje vseh latentnih faktorjev



Slika: Učenje vseh latentnih faktorjev hkrati.

Fokus raziskovalne naloge

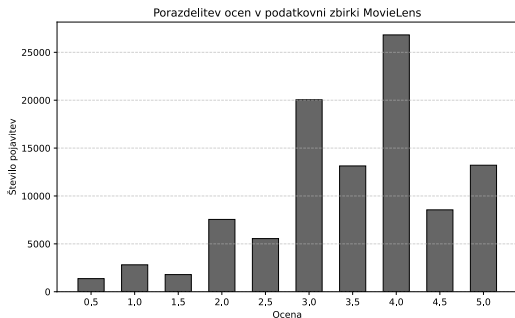
- ▶ Literatura nakazuje na **hitrejšo konvergenco učnega procesa pri pristopu sočasnega učenja**.
- ▶ Vse pomembnejša je **razložljivost modelov strojnega učenja**.
- ▶ Cilji raziskave:
 - ▶ Kateri pristop ima boljšo **napovedno točnost**?
 - ▶ Kateri pristop omogoča boljšo **razložljivost modela**?

Potek raziskave

1. Zbiranje literature o dosedanjih implementacijah obeh načinov implementacije matrične faktorizacije.
2. Pridobitev podatkovne zbirke, ki je bila uporabljena za analizo obeh pristopov.
3. Implementacija obeh pristopov v Python okolju z uporabo knjižnice NumPy.
4. Evalvacija modelov za napovedno napako in razložljivost latentnih faktorjev.

Porazdelitev ocen filmov

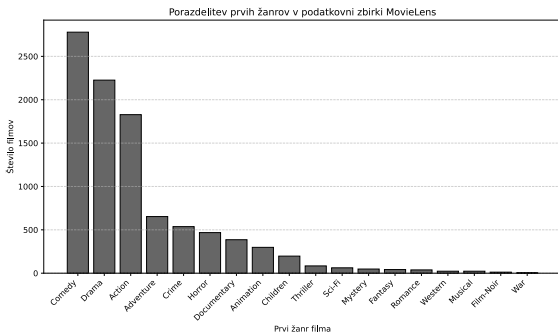
- ▶ Podatkovna zbirka MovieLens.
- ▶ Zbirka vsebuje 100.000 ocen za 10.000 filmov, (610 uporabnikov).
- ▶ Ocene večinoma srednje do visoke, kar kaže na pristranskost.



Slika: Porazdelitev ocen v podatkovni zbirki.

Porazdelitev filmskih žanrov

- ▶ Prevladovali so žanri komedija, drama in akcija.
- ▶ Redkejši žanri so bili bistveno slabše zastopani.
 - ▶ To je vplivalo na razložljivost modela.



Slika: Porazdelitev žanrov v podatkovni zbirki.

Implementacija

- ▶ Implementacija obeh pristopov matrične faktorizacije.
 - ▶ Python okolje, NumPy knjižnica.
- ▶ Za učenje uporabljen **gradientni spust**.
- ▶ Za oba pristopa uporabljene iste vrednosti hiperparametrov, kjer je bilo to mogoče.
- ▶ Za oba modela uporabljeni ista učna in validacijska množica.

Parameter	Postopno	Sočasno
Latentni faktorji	40	enako kot postopno
Ocene (učenje / validacija)	80668 / 20168	enako kot postopno
Epoch	100 na faktor	100
Hitrost učenja	0,001	enako kot postopno

Tabela: Primerjava uporabljenih hiperparametrov pri postopnem in sočasnem pristopu.

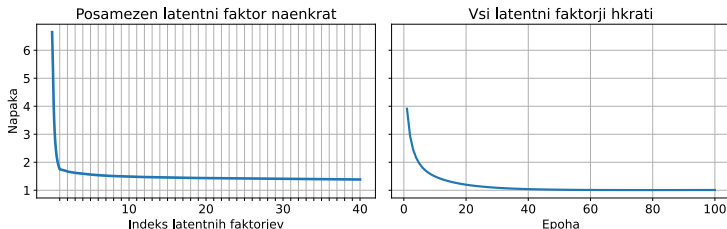
Evalvacija

- ▶ Napovedna točnost ocenjena z metrikami: **RMSE** (koren povprečne kvadratne napake).
- ▶ Razložljivost ocenjena z vizualizacijo in statističnimi testi.
- ▶ Vizualizacija: prva dva latentna faktorja filmskih vektorjev.
- ▶ Statistični testi:
 - ▶ **ANOVA** za preverjanje razlik med žanri.
 - ▶ **Parni t-testi** z Bonferronijevo korekcijo za primerjave med žanri.

Rezultati napovedne napake

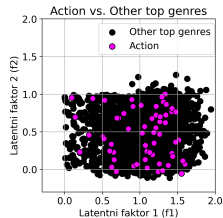
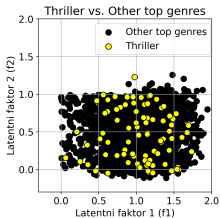
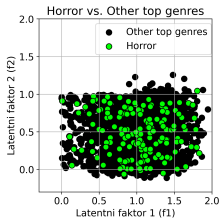
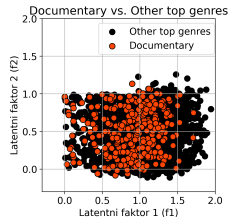
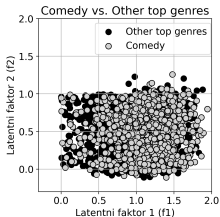
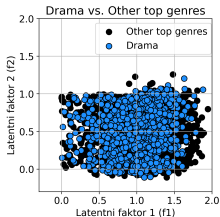
Metrika	Postopno	Sočasno
RMSE (validacija)	1,384	1,010
Čas treniranja	19 min	44 s

Tabela: Primerjava napovedne točnosti in časa treniranja med postopnim in sočasnim pristopom.



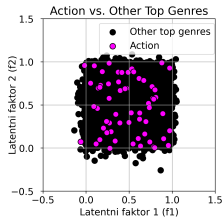
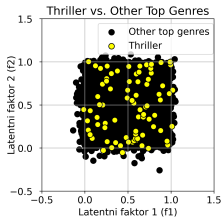
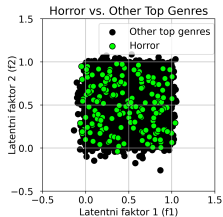
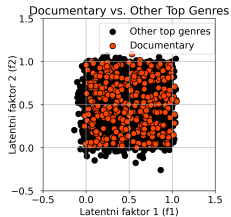
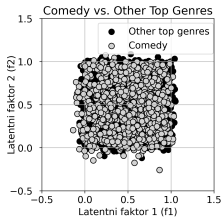
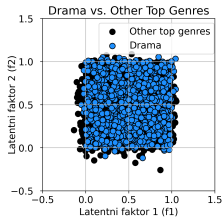
Slika: Validacijska napaka med treniranjem modelov.

Pristop postopnega treniranja latentnih faktorjev



Slika: Najpogostejših 6 žanrov v latenten prostoru prvih dveh faktorjev. ▶

Pristop sočasnega treniranja latentnih faktorjev



Slika: Najpogostejših 6 žanrov v latentnem prostoru prvih dveh faktorjev ▶

Analiza variance (ANOVA)

- ▶ Analiza preverja, ali obstajajo razlike med žanri glede na prvi latentni faktor.
- ▶ Statistično značilne razlike so bile zaznane le pri **postopnem pristopu**.
- ▶ Pri sočasnem pristopu ni bilo zaznanih značilnih razlik.

Pristop treniranja	p -vrednost
Postopno treniranje	$3,53 \times 10^{-17}$
Sočasno treniranje	0,2611

Tabela: Rezultati enosmerne ANOVA analize za prvi latentni faktor.

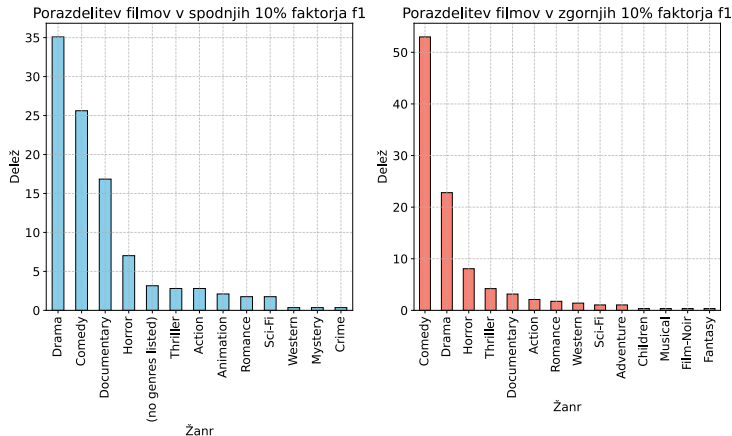
Parni t-testi z Bonferronijevo korekcijo

- ▶ Nadaljnja analiza razložljivosti je bila izvedena samo za **postopni model**.
- ▶ T-testi so pokazali pomembne razlike med nekaterimi pari žanrov.

Par žanrov	Prvoten p	Prilagojen p	Značilno
Comedy vs. Drama	0,0001	0,0001	da
Comedy vs. Documentary	0,0001	0,0001	da
Drama vs. Documentary	0,0001	0,0001	da
Documentary vs. Horror	0,0015	0,0232	da

Tabela: Rezultati parnih t-testov za postopno treniranje.

Žanrska sestava v prvem in zadnjem decilu



Slika: Žanrska sestava v prvem in zadnjem decilu prvega latentnega faktorja za pristop postopnega treniranja.

Primerjava pristopov učenja

- ▶ Rezultati kažejo na pomembne razlike med pristopoma učenja.
- ▶ Razlike izhajajo iz načina učenja latentnih faktorjev.

Sočasno učenje

- ▶ Vsi latentni faktorji se optimizirajo hkrati.
- ▶ Hitrejša konvergenca in nižja napovedna napaka.
- ▶ Sistem se optimizira kot celota.

Postopno učenje

Opis

- ▶ Vsak latentni faktor se uči ločeno.
- ▶ Ni motenj iz drugih faktorjev.
- ▶ Faktorji so bolj strukturirani glede na vsebino (npr. žanri).

Slabost

Višja napovedna napaka zaradi omejene sposobnosti učenja več vidikov hkrati.

Razložljivost in omejitve

- ▶ Le nekateri faktorji kažejo jasne razlike med žanri.
- ▶ Bolj razložljivi za pogoste žanre.
- ▶ Razložljivost je odvisna od porazdelitve podatkov.

Interakcije med faktorji

Pri sočasnem učenju lahko posamezni faktorji izgubijo razložljivost zaradi visoke medsebojne soodvisnosti.

Praktične implikacije

Sočasno učenje

- ▶ Učinkovito pri komercialnih aplikacijah.
- ▶ Cilj: nizka napovedna napaka.

Postopno učenje

- ▶ Primerno za aplikacije, kjer je pomembna razlaga (npr. izobraževanje, zdravstvo).
- ▶ Prispeva k večjemu zaupanju v sistem.

Možnosti za prihodnje delo

- ▶ Raziskati hibridne pristope.
- ▶ Uporaba regularizacije za spodbuditev razložljivosti.
- ▶ Metode za nadzorovano strukturiranje latentnega prostora.
- ▶ Večfazni modeli:
 - ▶ 1. faza: optimizacija napovedne natančnosti.
 - ▶ 2. faza: izboljšanje interpretabilnosti.

Literatura I

-  F. M. Harper and J. A. Konstan, “The movielens datasets: History and context,” *ACM Transactions on Interactive Intelligent Systems (TiiS)*, vol. 5, no. 4, pp. 19:1–19:19, 2015.
-  S. Funk, “Netflix update: Try this at home.”
<https://sifter.org/simon/journal/20061211.html>, 2006.
Accessed: 2025-05-18.
-  A. Falini, “A review on the selection criteria for the truncated svd in data science applications,” *Journal of Computational Mathematics and Data Science*, vol. 5, p. 100064, 2022.

Literatura II

-  G. Takács, I. Pilászy, B. Németh, and D. Tikk, “Scalable collaborative filtering approaches for large recommender systems,” *J. Mach. Learn. Res.*, vol. 10, p. 623–656, June 2009.
-  P. Linardatos, V. Papastefanopoulos, and S. Kotsiantis, “Explainable ai: A review of machine learning interpretability methods,” *Entropy*, vol. 23, no. 1, 2021.
-  G. Adomavicius and A. Tuzhilin, “Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions,” *IEEE Transactions on Knowledge and Data Engineering*, vol. 17, no. 6, pp. 734–749, 2005.