

Triplet Losses-based Matrix Factorization for Robust Recommendations

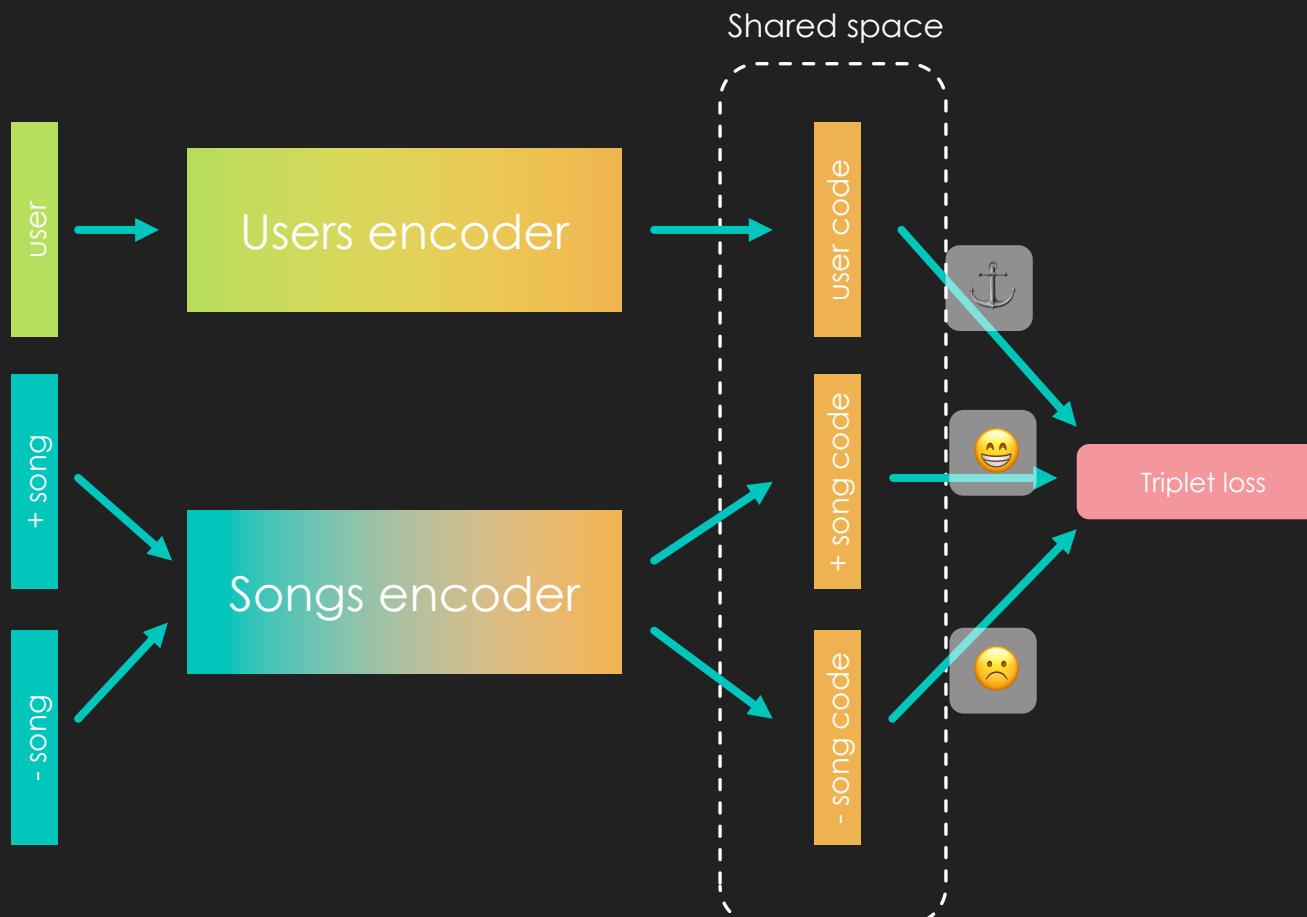
EvalRS @ CIKM 2022
October 21, 2022

Flavio Giobergia
Politecnico di Torino, Italy

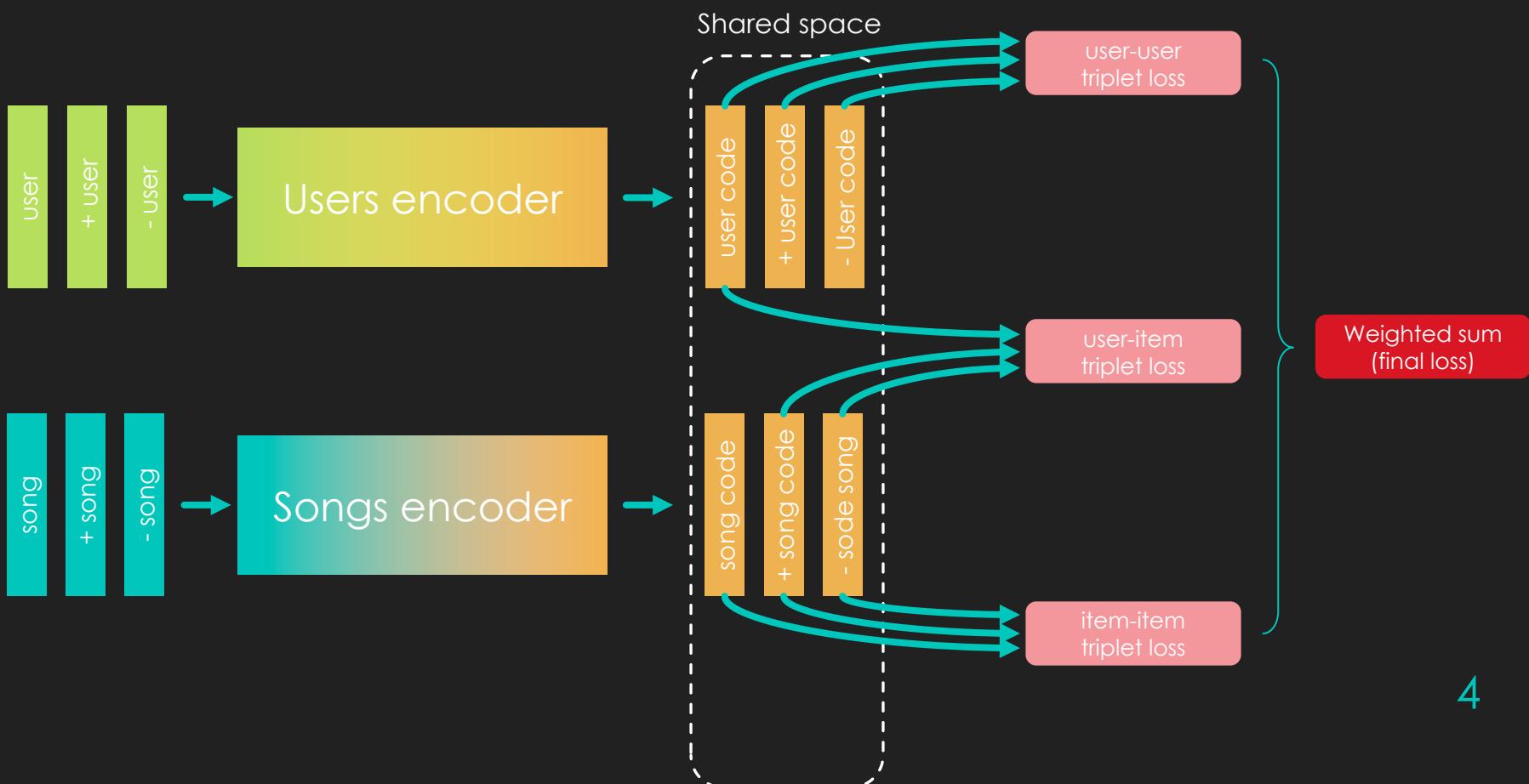
Challenge recap

- Songs recommendation
 - Based on users' previous interactions
- Multi-faceted evaluation (9 + 1 metrics!)
 - Standard, per group, behavioral
- Personal experience with recommender systems: ~ 0...
 - ... But how hard can they be?

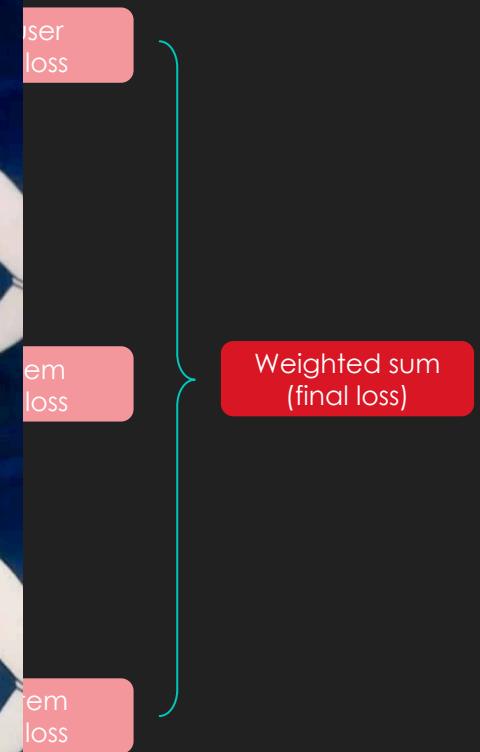
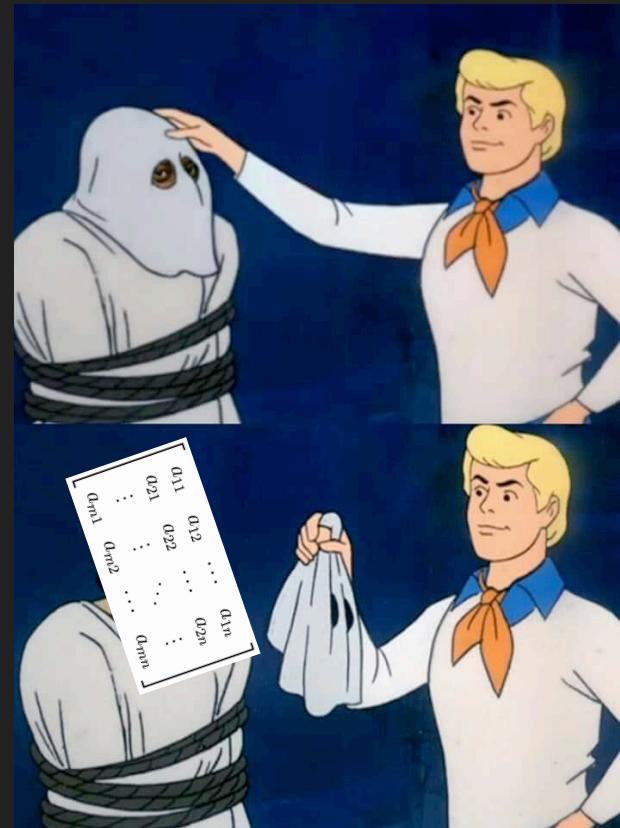
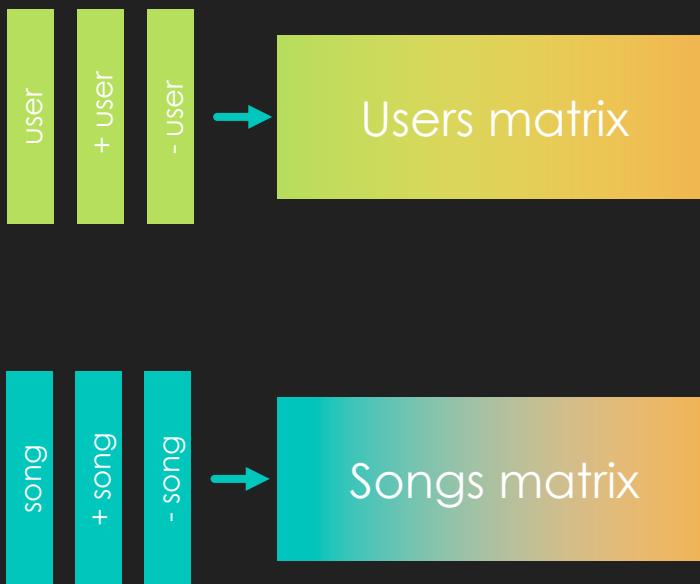
Solution v1 (triplet loss)



Solution v2 (triplet losses)



Solution v3 (keep it simple!)



Matrices initialization

- word2vec – based
 - One sentence for each “song”
 - Tokens for:
 - Users who listened to that song
 - Author of the song
 - Albums where the song shows up
 - Initialization w/o fine tuning works great! (hit rate-wise)

user=User1

user=User2

...

song=SongId

artist=ArtistId

album=Album1

album=Album2

...

w2v sentence

Samples weighting

- Each (user, song) pair is weighted differently
- Users and songs are assigned a “rarity” based on:
 - Age, gender, country, activity (for users)
 - Song and artist popularity (for songs)
- “Rarer” pairs will be weighted more in the loss

Additional metrics

○ Consistency

- “How well does the model perform across separate folds?”
- Mean (over all metrics) of the coefficient of variation across folds

The model is less consistent on some metrics!

Metric	Consistency
Hit rate	0.9579
MRR	0.8885
Country (MRED)	0.8057
User activity (MRED)	0.9312
Track popularity (MRED)	0.6938
Artist popularity (MRED)	0.7828
Gender (MRED)	0.4573
Being less wrong	0.9949
Latent diversity	0.9976
Overall	0.8344

○ Variance agreement

- Does the model match each user's preference in terms of variance?
 - e.g. in terms of artists
- Pearson correlation between Gini impurity of user's preferences and Gini impurity of model's recommendation for the user
- VA w.r.t. artists: ~ 0.25
 - (Random would be 0)

Numbers!

Solution	Score	Hit rate	MRR	Country (MRED)	User activity (MRED)	Track popularity (MRED)	Artist popularity (MRED)	Gender (MRED)	Being less wrong	Latent diversity
team#1	1.702570	0.015484	0.005859	-0.004070	-0.006932	-0.002044	-0.001688	-0.000956	0.424817	-0.121655
team#2	1.552977	0.016065	<u>0.001727</u>	-0.003727	-0.002913	-0.002307	-0.001047	-0.000692	0.363927	-0.296403
Proposed solution	1.330388	0.015565	0.001677	-0.004036	-0.003504	-0.004444	-0.000867	-0.000797	0.281863	-0.272944
team#4	1.184669	<u>0.021619</u>	<u>0.002044</u>	-0.005366	-0.004417	<u>-0.003191</u>	-0.001542	-0.001299	<u>0.320594</u>	-0.317348
team#5	1.138580	<u>0.018819</u>	0.001071	-0.005213	-0.005174	<u>-0.005043</u>	-0.001234	-0.002774	0.280727	<u>-0.244437</u>
team#6	1.028222	0.015006	0.001127	-0.005448	-0.007534	-0.005261	-0.001202	-0.003369	<u>0.316226</u>	-0.309870
team#7	0.752576	<u>0.017874</u>	0.001655	-0.006193	-0.010749	-0.004483	-0.002132	-0.004541	<u>0.322594</u>	-0.324841
team#8	0.429596	<u>0.018173</u>	0.001632	-0.005482	-0.011190	-0.007588	-0.002513	-0.004329	<u>0.322767</u>	-0.324794
team#9	-1.154413	<u>0.032359</u>	<u>0.002054</u>	-0.010624	-0.013020	-0.021992	-0.003122	-0.008458	<u>0.295921</u>	-0.330054
baseline	-1.212213	0.036343	<u>0.003694</u>	-0.009016	-0.022362	-0.011082	-0.007150	-0.006084	<u>0.375774</u>	-0.307977

:)



Flavio Giobergia

flavio.giobergia@polito.it

<https://github.com/fgiobergia/CIKM-evalRS-2022>

<https://arxiv.org/abs/2210.12098>