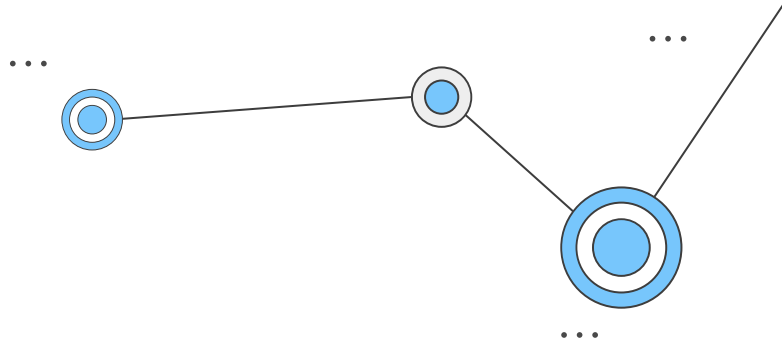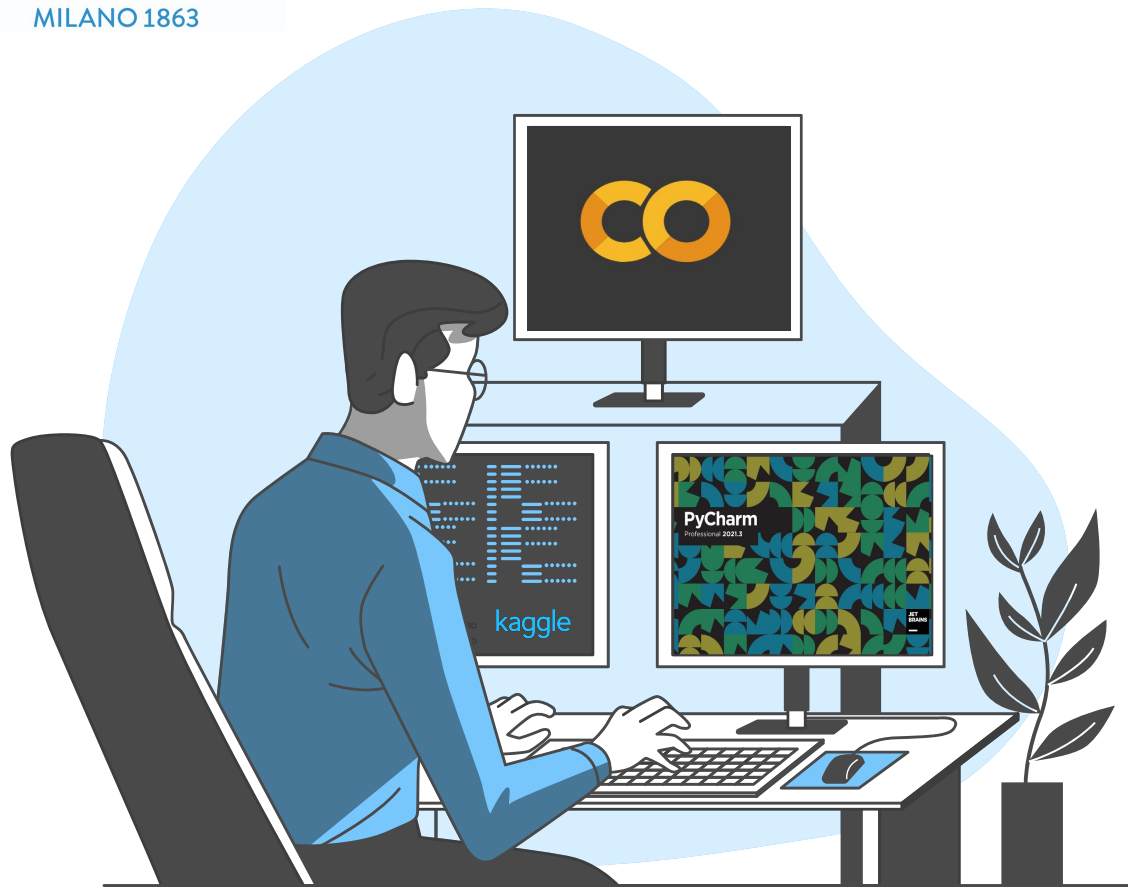# RecSys Challenge

a.y 2021-2022

Lorenzo Mainetti

Andrea Menta

# Timeline

**1** Data Exploration
Analysis of the URM and
ICM structures

**2** Profiling and Performance Analysis
Hyperparameters optimization and
models comparison

**3** Hybrid Experiments
Various combinations of the base models
with different techniques

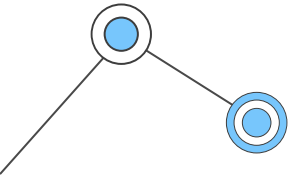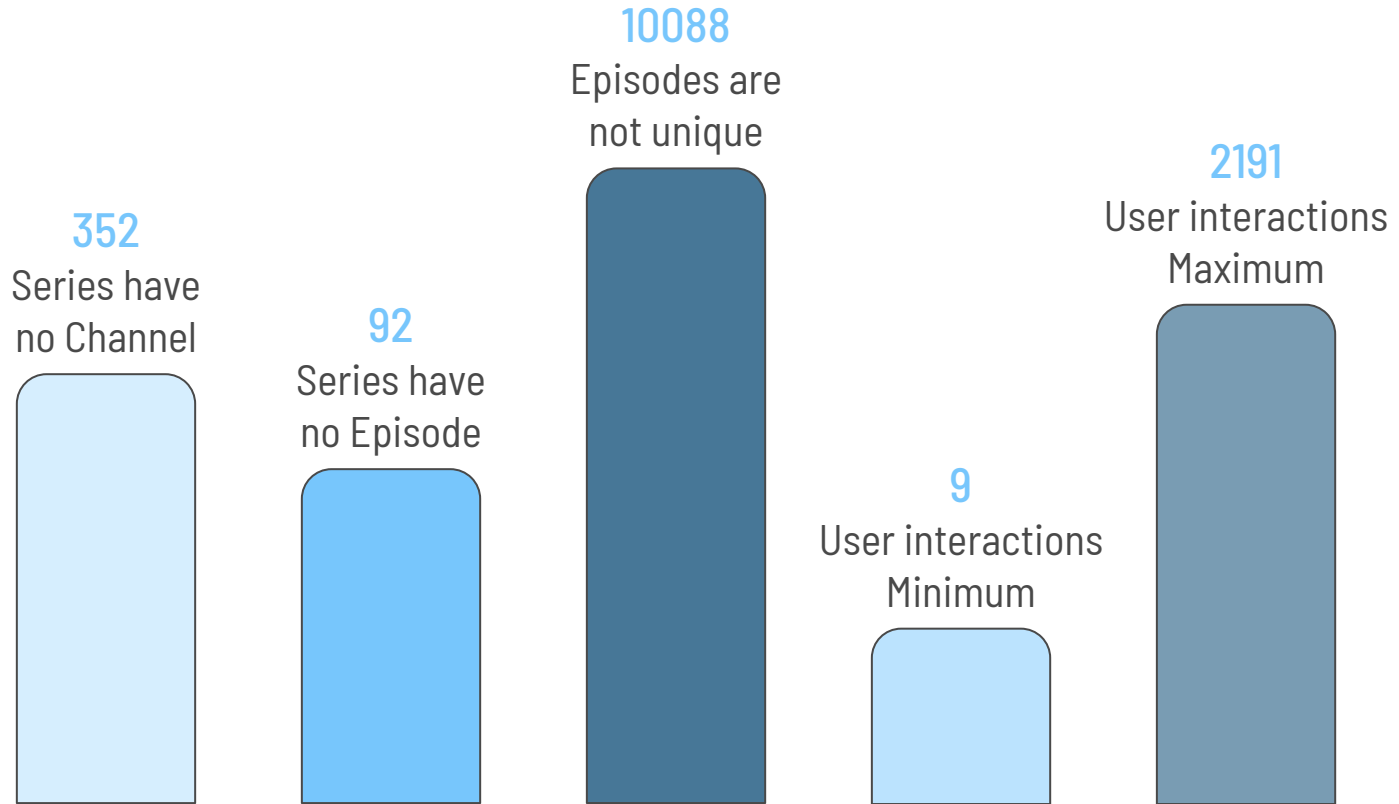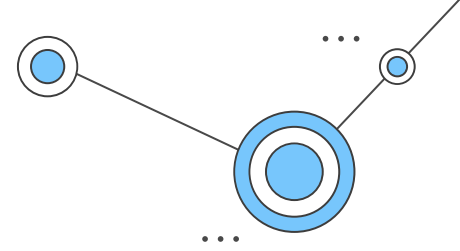**4** Fine Tuning
Best model improvements

# 1 – Data Exploration

**352**
Series have
no Channel

**92**
Series have
no Episode

**10088**
Episodes are
not unique

**9**
User interactions
Minimum

**2191**
User interactions
Maximum

# 2.1 – Profiling

| | MAP@10 | Precision | Recall | Novelty |
|---|---|---|---|---|
| IALS | 0.2338 | 0.3824 | 0.0679 | 0.0056 |
| PureSVD | 0.2340 | 0.3814 | 0.0668 | 0.0055 |
| RP3Beta | 0.2263 | 0.3824 | 0.0679 | 0.0056 |
| SLIM BPR | 0.2359 | 0.3830 | 0.0691 | 0.0055 |
| SLIM ElasticNet | 0.2501 | 0.4039 | 0.0730 | 0.0055 |
| UserItemKNNCF | 0.2345 | 0.3809 | 0.0669 | 0.0054 |

\* metrics evaluated on the validation set with a split 80/20
\*\*base models not in the table were not effective and were not used in our final model


EASE_R


EASE_R


EASE_R
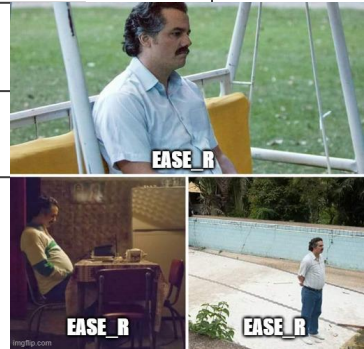
# 2.2 – Performance Analysis
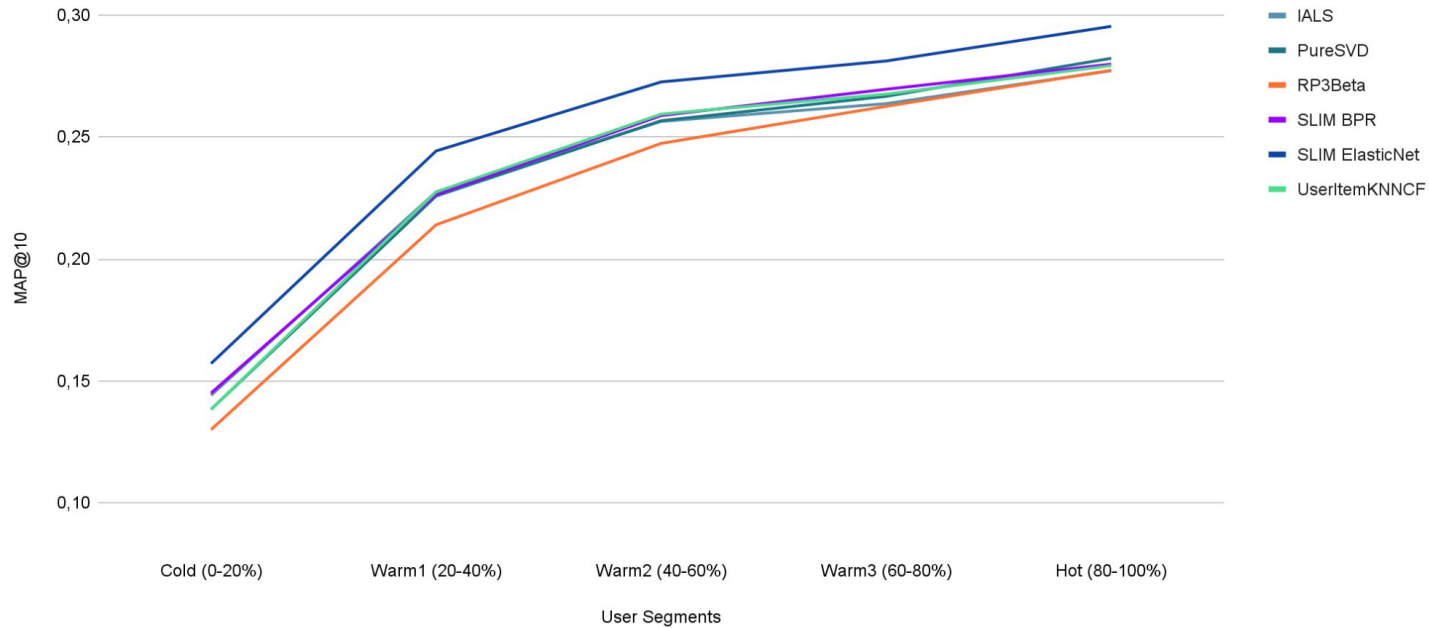
# 3 – Hybrid Experiments

### Linear Combination

Weighted sum of predictions

### Similarity Combination

Weighted sum of similarity matrices (merge models with similar structure)

### Pipelining

Output ratings of one algorithm is fed as input to the next algorithm

### User Segmentation

Split users based on number of interactions

### Stacking

Add context information by stacking URM and ICM

### Co-training

Simultaneous training of models

### Ensemble

Multiple trainings of different splits of the dataset

# 4.1 – What made the difference

## Hierarchical structure
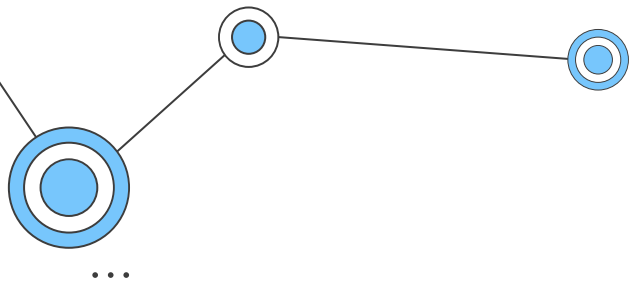Nested hybrids to increasingly improve performance

## Coupled hybrids
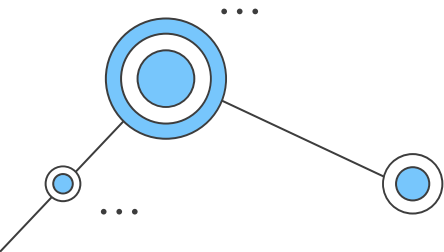Hybrid made of just two submodels

## Focus on weak points
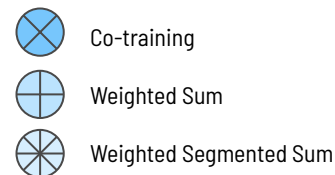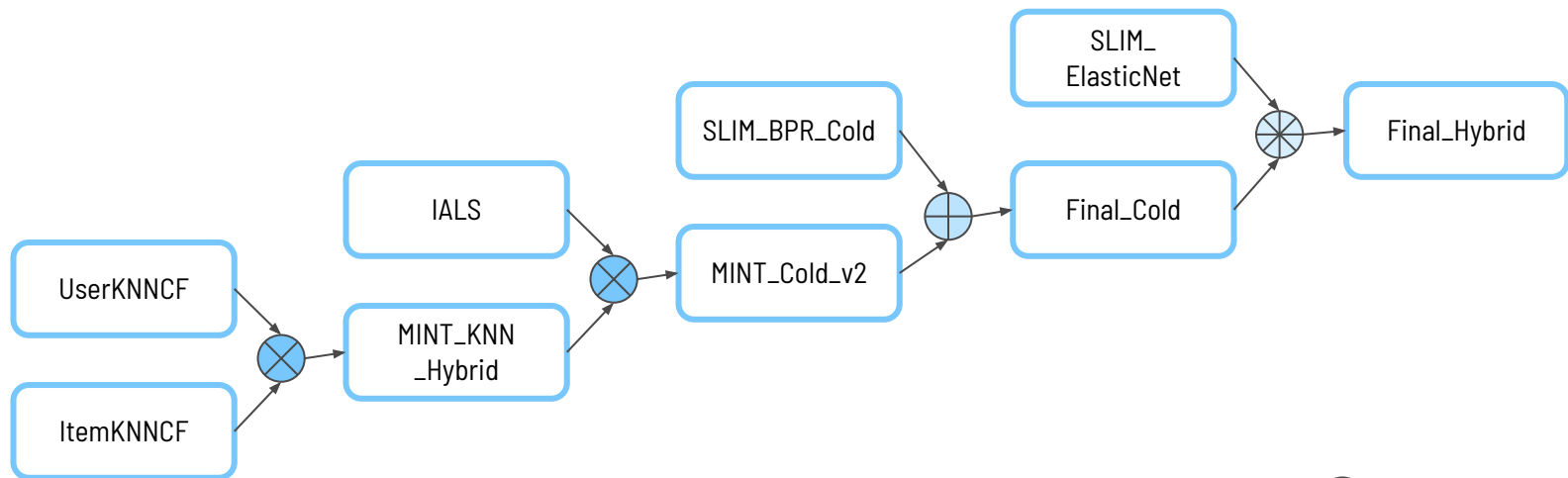Some models were trained on the cold user segment

*"Dacrema's words echoed...
There's an hyperparameter for
every model, but you must find it"*

—Maurizio Ferrari Dacrema

# 4.2 – Best Model

Private leaderboard
MAP score: **0.50787**



UserKNNCF

ItemKNNCF

IALS

MINT_KNN_Hybrid

SLIM_BPR_Cold

MINT_Cold_v2

SLIM_ElasticNet

Final_Cold

Final_Hybrid

Co-training

Weighted Sum

Weighted Segmented Sum

# Thanks!

Do you have any questions?

lorenzo.mainetti@mail.polimi.it
andrea.menta@mail.polimi.it

Code available on GitHub