ML algorithms for product recommendations

Sergei Bulaev

Senior Data Scientist

Epam Systems

June 2023

About me

- Senior Data Scientist at EPAM Systems
- 5+ years in Data Science
- MSc Master of Data Science

Agenda

1. Current challenges
2. Potentional solution
3. Benchmarks on H&M dataset

Current challenges

Challenges. Price of a wrong prediction

Challenges. Price of a wrong prediction

Challenges. **Price of a wrong prediction**

We are in a risk-free zone

Challenges. **Goals**

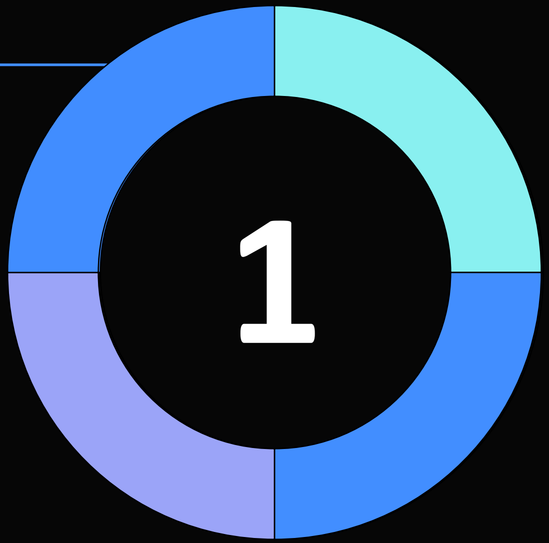
<A Review of Modern Fashion Recommender Systems>

1. Outfit generation
2. Outfit recommendation
3. Outfit compatibility prediction
4. Fill In the blank
5. Pair recommendation

* https://arxiv.org/pdf/2207.02757.pdf

Challenges. Modeling

Reproducibility of results



Reproducibility of results

<Are We Really Making Much Progress?>*

Conference	Rep. ratio	Reproducible
KDD	3/4 (75%)	[17], [23], [48]
RecSys	1/7 (14%)	[53]
SIGIR	1/3 (30%)	[10]
WWW	2/4 (50%)	[14], [24]

Total: 7/18 (39%)

* https://arxiv.org/pdf/2211.01261.pdf

Challenges. Modeling

Reproducibility of results

Absence of a strong baseline

2

Absence of a strong baseline

< Revising the Performance of IALS on the Item Recommendation Benchmarks >*

Dataset	Method	Recall@20
ML20M	RecVAE [25]	0.414
	H+Vamp (Gated) [14]	0.413
	RaCT [18]	0.403
	Mult-VAE [17]	0.395
	LambdaNet [4]	0.395
	iALS	**0.395**
	EASE [26]	0.391
	CDAE [28]	0.391
	Mult-DAE [17]	0.387
	SLIM [19]	0.370
	iALS	***0.363***
	iALS	***0.360***
	WARP [27]	0.314
	Popularity	0.162

* https://arxiv.org/pdf/2110.14037.pdf

Challenges. Modeling

Reproducibility of results

Absence of a strong baseline

3

Corr(online, offline) metrics

Corr(online, offline) metrics

< Off-line vs. On-line Evaluation of Recommender Systems in Small E-commerce >*

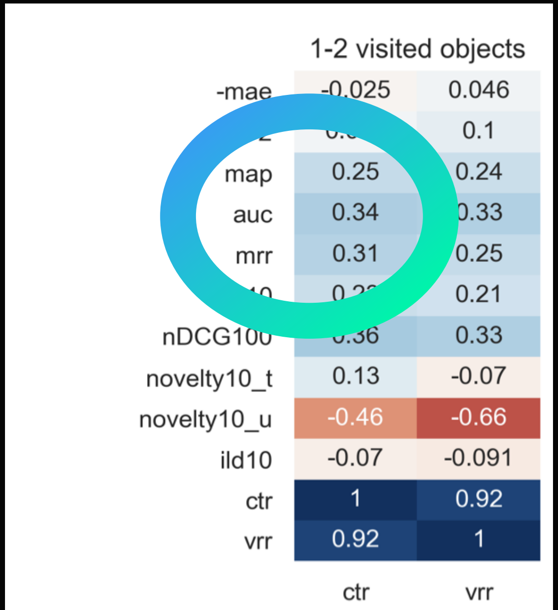


Figure 3: Spearman's correlation between off-line and on-line evaluation metrics for various user model sizes.

* https://arxiv.org/pdf/1809.03186.pdf

Challenges. Modeling

Reproducibility of results

Absence of a strong baseline

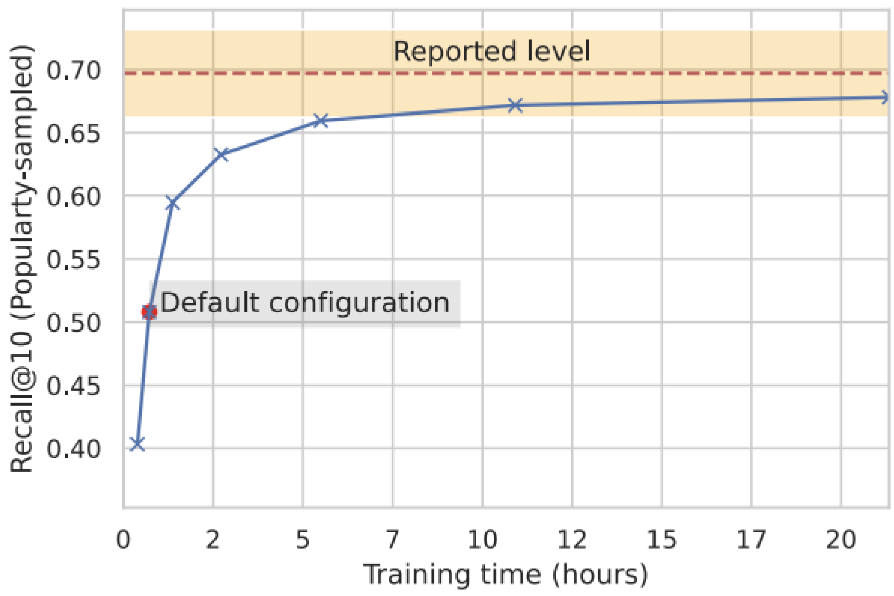
4

Long retraining time

Corr(online, offline) metrics

Long retraining time

< A Systematic Review and Replicability Study of BERT4Rec for Sequential Recommendation >*



(a) Effect on popularity-sampled Recall@10

* https://arxiv.org/pdf/2207.07483.pdf

Solution

Basic recommendation system

```
df.groupby(product_id)[sold_units].sum()
```

Usually considered as a baseline in scientific articles

Basic recommendation system

- feature_list = [product_id, gender, age, region, segment]
  df.groupby(feature_list)[sold_units].sum()

Basic recommendation system

- df = df filter by 2 previous weeks
  feature_list = [product_id, gender, age, region, segment]
  df.groupby(feature_list)[sold_units].sum()

Basic recommendation system

Customer put the product into the basket.

- **Filter users who bought the current product**

df = df filter by 2 previous weeks

feature_list = [product_id, gender, age, region, segment]

df.groupby(feature_list)[sold_units].sum()

Basic recommendation system

- **Mix recommendations with repurchase ratio**

  Filter users who bought the current product

  df = df filter by 2 previous weeks

  feature_list = [product_id, gender, age, region, segment]

  df.groupby(feature_list)[sold_units].sum()

Basic recommendation system

Personalised Popular Product

Collaborative filtering

Quadratic loss function

$$L(W, H) = L_S(W, H) + L_I(W, H) + R(W, H)$$

Basic recommendation system

$$L(W, H) = L_S(W, H) + L_I(W, H) + R(W, H)$$

$$\sum_{(u,i)\in S} \left(\hat{y}\left(y, i\right) - 1\right)^2$$

$$\alpha_0 \sum_{u\in U} \sum_{i\in I} (\hat{y}(u, i))^2$$

$$\lambda\left(\sum_{u\in U} \#\frac{v}{u} \|w_u\|^2 + \sum_{i\in I} \#\frac{v}{i} \|h_i\|^2\right)$$

Deploy it
&
Generate profit

Thank you!

Sergei Bulaev

Senior Data Scientist

But there is no

AI (Artificial Intelligence)
Neural Networks
Deep Learning
Machine learning

LightGBM

Yandex Catboost

Advanced recommendation system

Features generation

- Users behaviour
- Items sales performance
- Items similarity
- Aggregative statistics

Candidate generation

- Personalised Popular Product
- Collaborative filtering

- Build the process
- Gain an experience
- Build relationships
- Gather the right data
- Gain Trust
- Cover most of business lines
- Generating profit

Enjoy the research

Advanced recommendation system

NFM

MultiVAE

GCN

GRU4Rec

DiffuRec

SLIM

SHAN

SASREC

ChatGPT

xDeepFM

LightGCN

BERT4Rec

Comparison

H&M recommendation system challenge dataset

We are given a dataset consisting of 3 tables and a folder with pictures.

Customers	Products	Transactions	Pictures of H&M assortment
1 371 980 unique customers	**105 542** unique products	**31 788 324** transactions	**30 GB** of pictures

Target:
To predict what articles each customer will purchase in the 7-day period

Evaluation Metric:
Mean Average Precision MAP@12

Key results

Applied models for recommendation systems

Recommendation algorithms

Popularity based

- Most popular (year, month, season, week, age)
- Pairing purchaised items
- Repurchase products by customers

MAP@12
0.027

Collaborative filtering

- ALS
- SVD
- FunkySVD
- FM

MAP@12
0.024

Hybrid models

- LightFM
- ALS with features

MAP@12
0.021

Graph based

- GCN
- LightGCN
- Fi-GNN
- GRU4Rec
- SHAN

MAP@12
0.028

Tree based

- LightGBM
- CatBoost

MAP@12
0.032

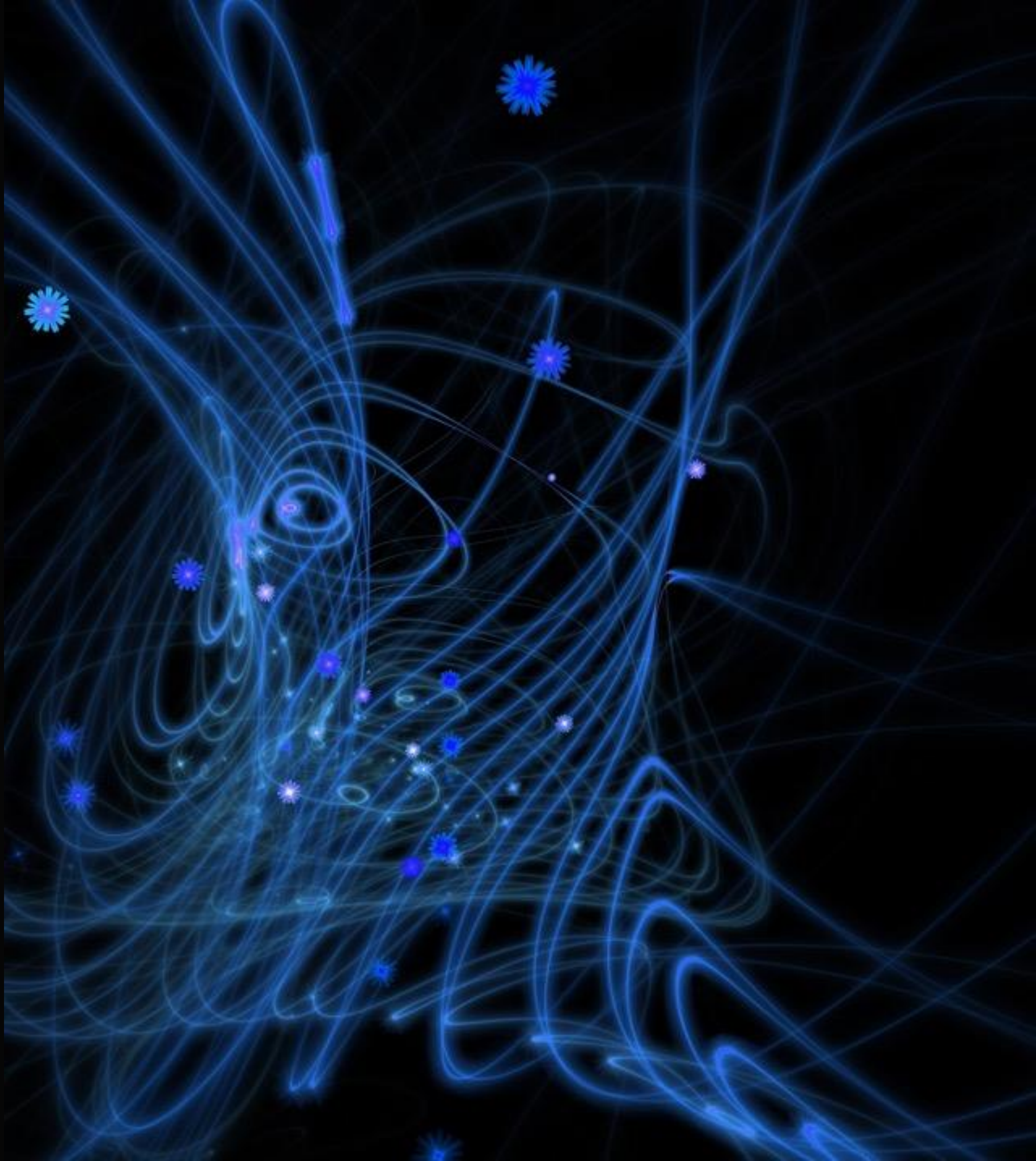
Key TakeAway

Simple and Fast Business results,
Rather than
State Of The Art Models

Thank you!

Sergei Bulaev

Senior Data Scientist



<epam>