Financial Asset Recommendation

Comparing the Performance of collaborative filtering, contextual and non-contextual bandit algorithms

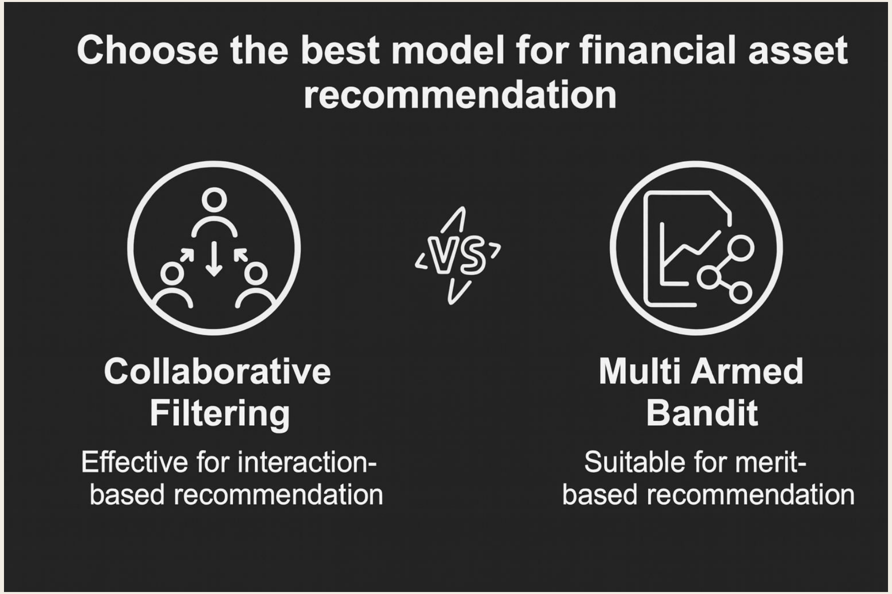
By: Bharat Khandelwal

AGENDA

- Introduction
- Collaborative Filtering
- Why Multi Armed Bandit?
- Contextual Multi Armed Bandit
- Dataset
- Experimental Setup
- Results
- Conclusion

INTRODUCTION

- A data-driven system that suggests optimal financial assets (e.g., stocks, funds, ETFs, bonds) for individual investors.
- The most common way of recommending a financial asset is through traditional risk-based portfolio allocation, where assets are suggested based on an investor's risk profile using rule-based or mean-variance (Markowitz) portfolio optimization methods.
- **Impact of FAR**:
  - Better Portfolio Performance
  - Customer Satisfaction
  - Minimizes emotional decision-making

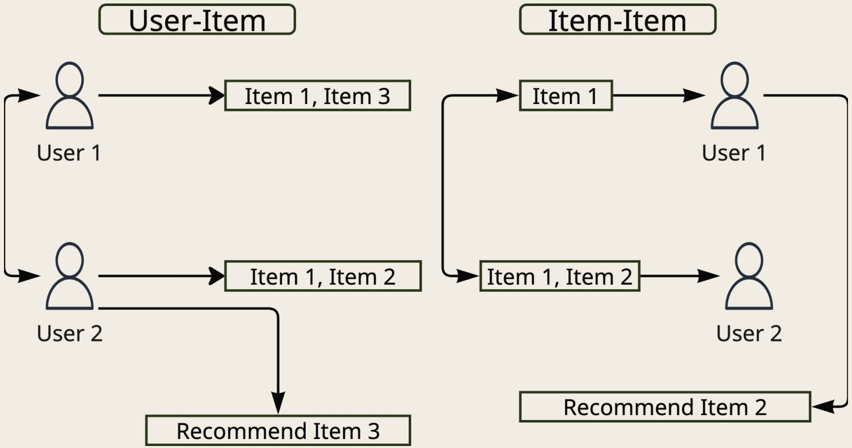


CHALLENGES

- Market volatility makes asset returns highly unpredictable.
- Hard to get real-world data
- Investor preferences and risk profiles vary widely and change over time.
- Sparse or limited historical interaction data reduces recommendation accuracy.
- Different assets have different liquidity and risk characteristics, making comparison difficult.
- Feedback is delayed because true reward is known only after days/weeks.

Collaborative Filtering (CF)

- Recommends assets by finding similar users based on shared purchase behavior.
- Uses a user-asset interaction matrix and similarity measures (e.g., cosine similarity).
- Assumes users with similar history will like similar assets in the future.



"Collaborative Filtering recommends assets by finding either similar users (User–Item CF) or similar assets (Item–Item CF) based on shared interaction patterns."

Why Multi Armed Bandit

- Provides a framework for sequential decision-making.
- Offers a principled way to evaluate merit-based recommendations by linking each asset choice to a measurable reward.
- Enables exploration vs. exploitation trade-offs, important when asset performance is uncertain.
- Allows testing whether contextual features (e.g., momentum, volatility) meaningfully relate to returns.
- Extends FAR research by introducing adaptive decision models not evaluated in prior benchmarks.

Contextual Multi Armed Bandit

- Contextual Multi-Armed Bandits extend standard bandits by using additional information (context), such as asset features to guide which arm (asset) to choose each round.

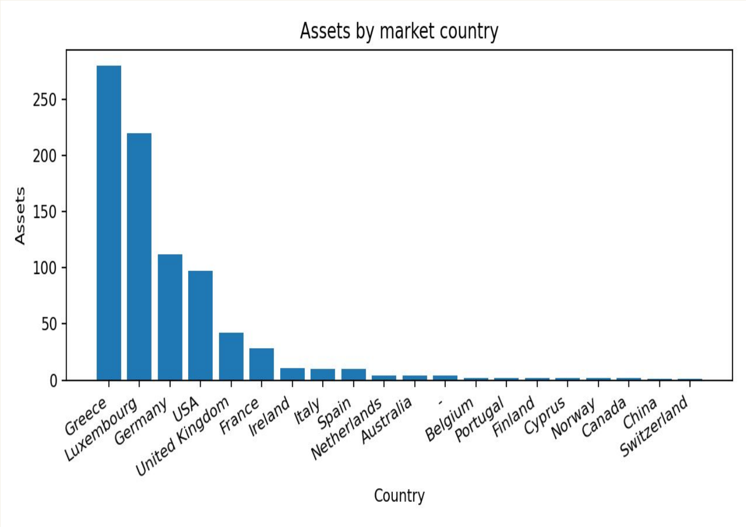
- LinUCB is a widely used contextual bandit algorithm that models the relationship between context features and expected reward using a linear function with confidence bounds.

Algorithm 1 LinUCB with disjoint linear models.

Inputs: $\alpha \in \mathbb{R}_+$
for $t = 1, 2, 3, \ldots, T$ **do**
  Observe features of all arms $a \in \mathcal{A}_t$: $\mathbf{x}_{t,a} \in \mathbb{R}^d$
  for all $a \in \mathcal{A}_t$ **do**
    if a is new **then**
      $\mathbf{A}_a \leftarrow \mathbf{I}_d$ (d-dimensional identity matrix)
      $\mathbf{b}_a \leftarrow \mathbf{0}_{d\times 1}$ (d-dimensional zero vector)
    end if
    $\hat{\boldsymbol{\theta}}_a \leftarrow \mathbf{A}_a^{-1}\mathbf{b}_a$
    $p_{t,a} \leftarrow \hat{\boldsymbol{\theta}}_a^\top \mathbf{x}_{t,a} + \alpha\sqrt{\mathbf{x}_{t,a}^\top \mathbf{A}_a^{-1}\mathbf{x}_{t,a}}$
  end for
  Choose arm $a_t = \arg\max_{a\in\mathcal{A}_t} p_{t,a}$ with ties broken arbitrarily, and observe a real-valued payoff r_t
  $\mathbf{A}_{a_t} \leftarrow \mathbf{A}_{a_t} + \mathbf{x}_{t,a_t}\mathbf{x}_{t,a_t}^\top$
  $\mathbf{b}_{a_t} \leftarrow \mathbf{b}_{a_t} + r_t\mathbf{x}_{t,a_t}$
end for

DATASET

- The FAR-Trans dataset (Sanz-Cruzado et al., 2024), containing real retail investor transactions, asset metadata, and market price histories from 2018–2022.
- Includes multiple components: customer profiles, asset information, transaction records, and daily close prices across different markets.
- Contains 806 unique assets, of which 321 have active investments, spread across 38 different markets.
- Records 703,303 price data points, enabling construction of weekly returns and market-driven features.



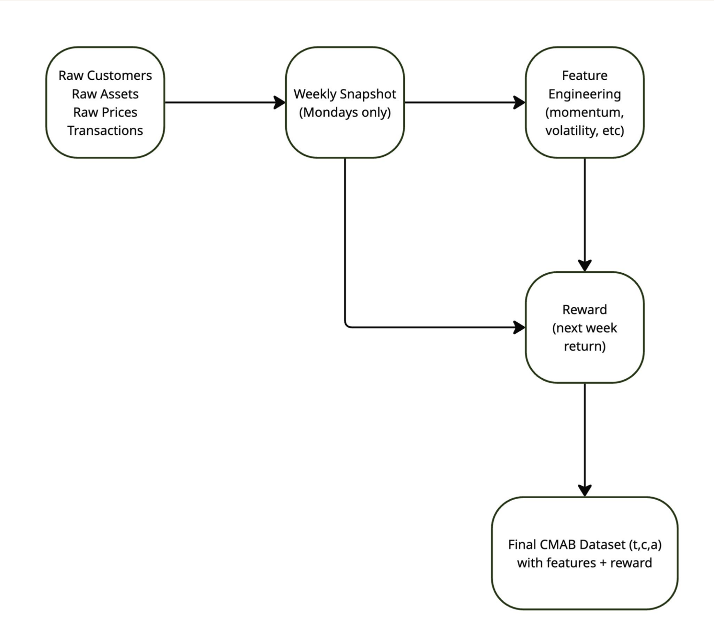
MAB Data STRUCTURE

Assets (Arms): All available assets at week *t*
Context:

- Asset momentum
- Asset volatility
- MarketID, Country
- Optional: customer risk

Action: Algorithm chooses one asset to recommend per timestamp
Reward:

- Next-week return of the chosen asset
- Used to update algorithm parameters



EXPERIMENTAL SETUP

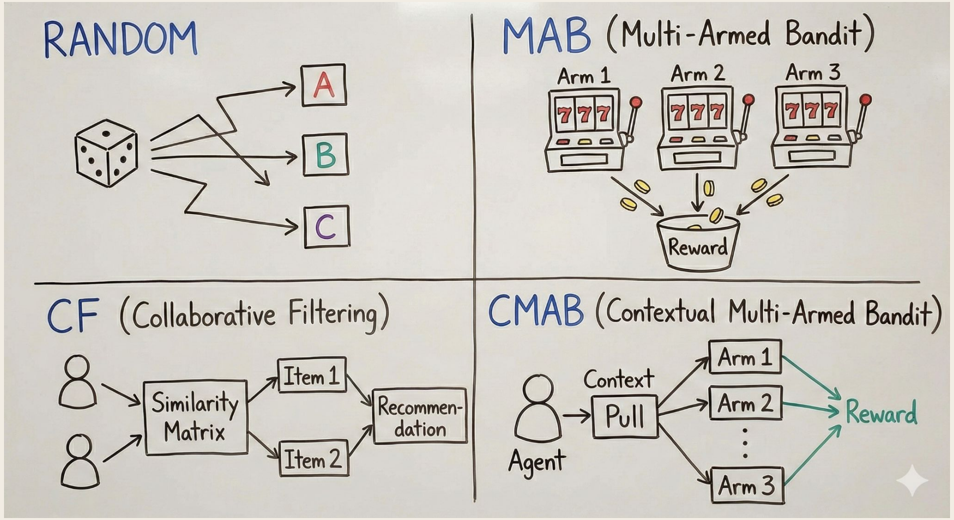
Experiments

- Ran two main scenarios:
    1. 68 assets × 3000/500 customers
    2. 253 assets × 1000/500 customers
- Sequential rounds: 253 weekly decisions

Models Compared

- Random baseline
- Item–Item Collaborative Filtering
- UCB1 (non-contextual bandit)
- LinUCB (contextual bandit)
- LinUCB with risk-level context included

Objective: Compare interaction-based vs feature-based vs adaptive bandit models.



EVALUATION METRICS

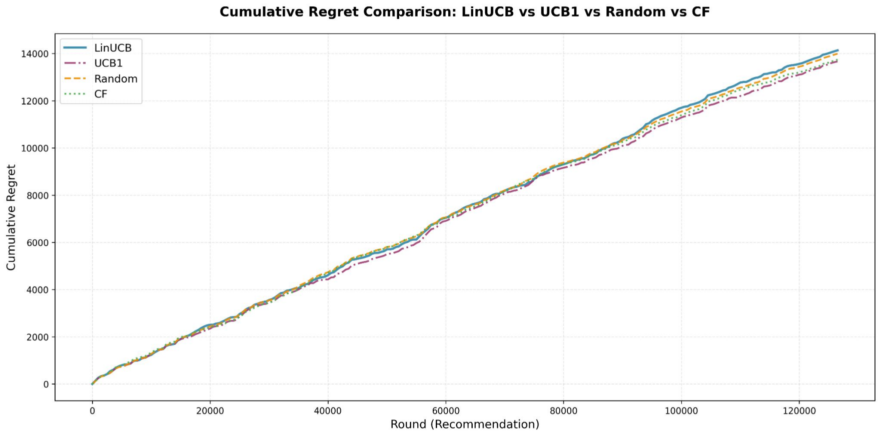
Primary Metric: Regret

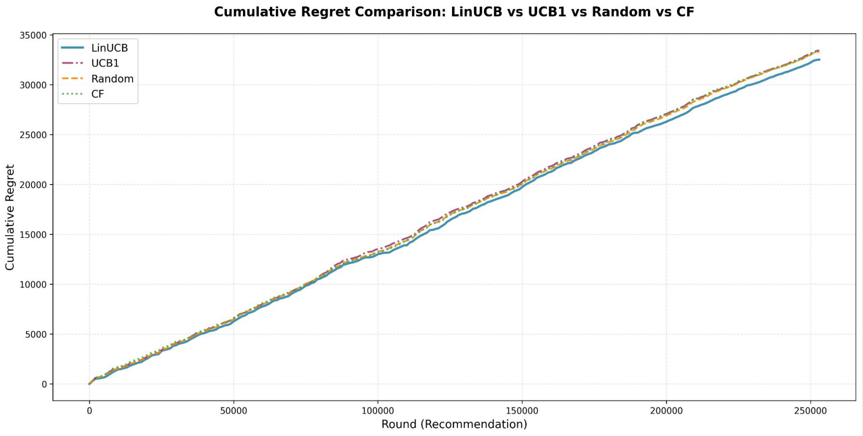
- Measures how much worse the algorithm performs compared to a perfect oracle that always picks the best asset each round.
- Lower regret = better adaptability & decision-making.

Why Regret?

- Suitable for sequential financial decisions.
- Captures the cost of exploration vs exploitation.
- Standard for benchmarking bandit algorithms.

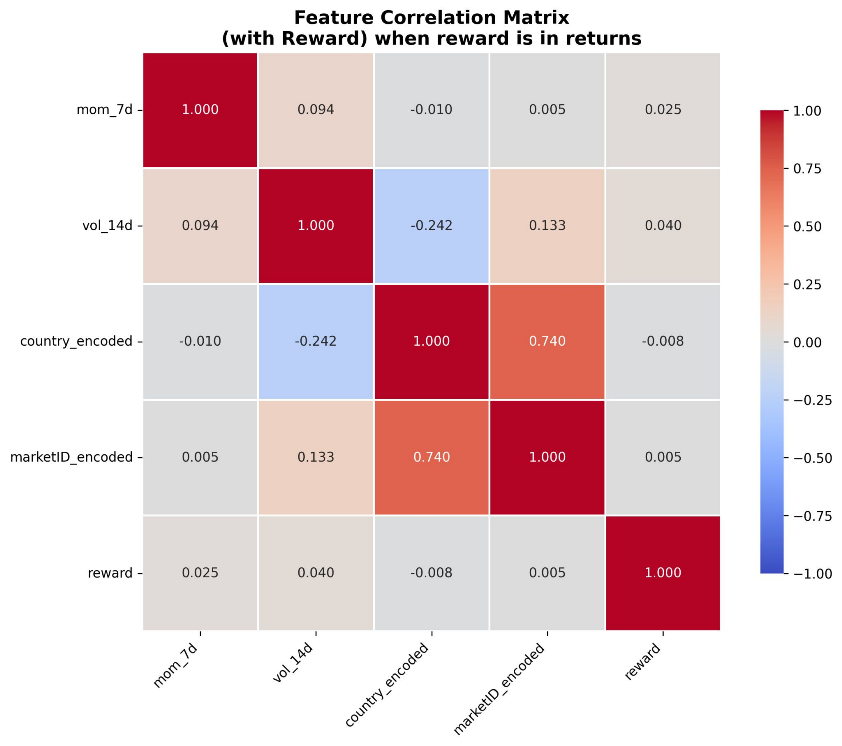
RESULTS





Scenario	Random	CF	UCB1	LinUCB	LinUCB (Risk)
1	13991.18	13743.74	13664.14	14134.75	–
2	84052.80	85153.74	82649.09	83347.65	–
3	16682.10	16220.62	16726.83	16262.33	–
4	33332.86	33404.51	33453.66	32524.67	–
5	33332.86	33404.59	33453.66	–	32524.67

- Cumulative regret grows almost linearly across all scenarios, showing that weekly asset returns are highly noisy and difficult to predict as major events like Covid-19 and war took place.
- LinUCB and UCB1 perform nearly identically, reflecting the weak correlation between contextual features and forward returns.
- Collaborative Filtering and Random baselines incur higher regret, confirming that static or interaction-based recommenders struggle in volatile financial environments.
- Engineered features (momentum, volatility, metadata) show minimal predictive power, limiting the advantage of contextual bandit methods.
- Including customer risk level does not improve performance, indicating that short-term returns are driven primarily by market behavior rather than user attributes.

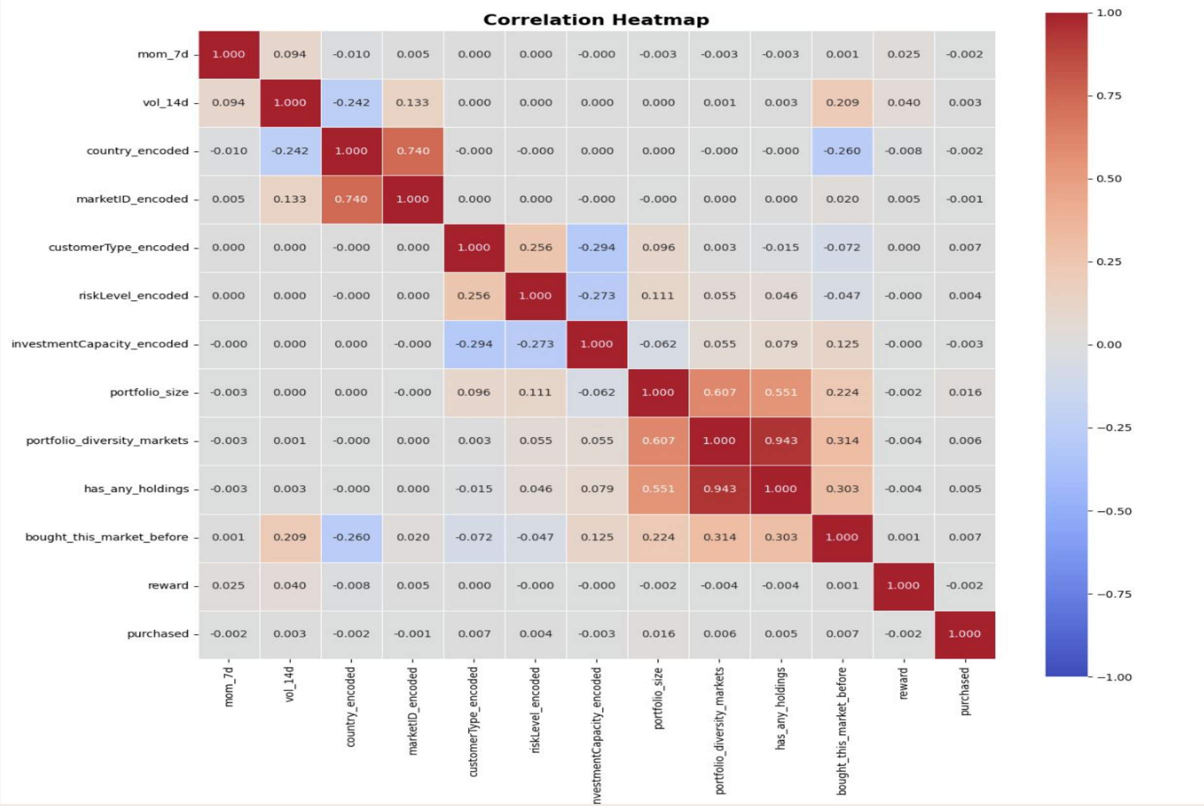


CONCLUSION

- Bandit algorithms provide a structured framework for studying sequential asset recommendation, but their effectiveness depends heavily on the predictive strength of available features.
- In the FAR-Trans setting, LinUCB and UCB1 show similar performance because short-term features (momentum, volatility) exhibit extremely weak correlation with future returns and non-contextual bandits do not perform well when the reward is temporal and non-stationary.
- Collaborative Filtering and Random baselines underperform, highlighting the limitations of static recommenders in volatile, non-stationary financial markets.
- Adding customer risk level does not improve outcomes, indicating that market-driven factors, rather than user characteristics, dominate short-term asset behavior.
- Overall, the results show that contextual bandits remain promising when informative signals exist, and regret is a valuable metric for evaluating learning behavior in uncertain financial environments.

FEEDBACK!

Appendix


Correlation Heatmap
mom_7d 1.000 0.094 -0.010 0.005 0.000 0.000 -0.000 -0.003 -0.003 -0.003 0.001 0.025 -0.002
vol_14d 0.094 1.000 -0.242 0.133 0.000 0.000 0.000 0.000 0.001 0.003 0.209 0.040 0.003
country_encoded -0.010 -0.242 1.000 0.740 -0.000 -0.000 0.000 0.000 -0.000 -0.000 -0.260 -0.008 -0.002
marketID_encoded 0.005 0.133 0.740 1.000 0.000 0.000 -0.000 -0.000 0.000 0.000 0.020 0.005 -0.001
customerType_encoded 0.000 0.000 -0.000 0.000 1.000 0.256 -0.294 0.096 0.003 -0.015 -0.072 0.000 0.007
riskLevel_encoded 0.000 0.000 -0.000 0.000 0.256 1.000 -0.273 0.111 0.055 0.046 -0.047 -0.000 0.004
investmentCapacity_encoded -0.000 0.000 0.000 -0.000 -0.294 -0.273 1.000 -0.062 0.055 0.079 0.125 -0.000 -0.003
portfolio_size -0.003 0.000 0.000 -0.000 0.096 0.111 -0.062 1.000 0.607 0.551 0.224 -0.002 0.016
portfolio_diversity_markets -0.003 0.001 -0.000 0.000 0.003 0.055 0.055 0.607 1.000 0.943 0.314 -0.004 0.006
has_any_holdings -0.003 0.003 -0.000 0.000 -0.015 0.046 0.079 0.551 0.943 1.000 0.303 -0.004 0.005
bought_this_market_before 0.001 0.209 -0.260 0.020 -0.072 -0.047 0.125 0.224 0.314 0.303 1.000 0.001 0.007
reward 0.025 0.040 -0.008 0.005 0.000 -0.000 -0.000 -0.002 -0.004 -0.004 0.001 1.000 -0.002
purchased -0.002 0.003 -0.002 -0.001 0.007 0.004 -0.003 0.016 0.006 0.005 0.007 -0.002 1.000
mom_7d
vol_14d
country_encoded
marketID_encoded
customerType_encoded
riskLevel_encoded
nvestmentCapacity_encoded
portfolio_size
portfolio_diversity_markets
has_any_holdings
bought_this_market_before
reward
purchased
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00

Appendix

- Contextual bandits only improve performance when contextual features are strongly predictive of reward. When context-reward correlation is low, contextual and non-contextual methods perform similarly. (“*A Survey on Practical Applications of Multi-Armed and Contextual Bandits*”, 2019) – **D.Bouneffouf & I.Rish**
- Volatility regimes and non-stationary drifts pose a challenge for contextual bandits, as the reward distribution changes faster than the algorithm can adapt.(“*Hedging using reinforcement learning: Contextual k-armed bandit versus Q-learning”*, 2023) – **L.Cannelli**
- Non-stationarity violates the stochastic assumptions of classical bandits, leading to poor regret performance.(“*Introduction to Multi-Armed Bandits*”, 2019) – **A.Slivkins**
- Shows that user preferences often change over time and under such non-stationarity, classical CF fails(“*Learning User Preferences in Non-Stationary Environments*”, 2021) – **Huleihel, Pal & Shayevitz**