University of Glasgow

FinTech Scotland

How to provide effective financial asset recommendations?

What's my next Investment? Automated Recommendations for Investors

05-10-2023 – Scottish Fintech Festival 2023

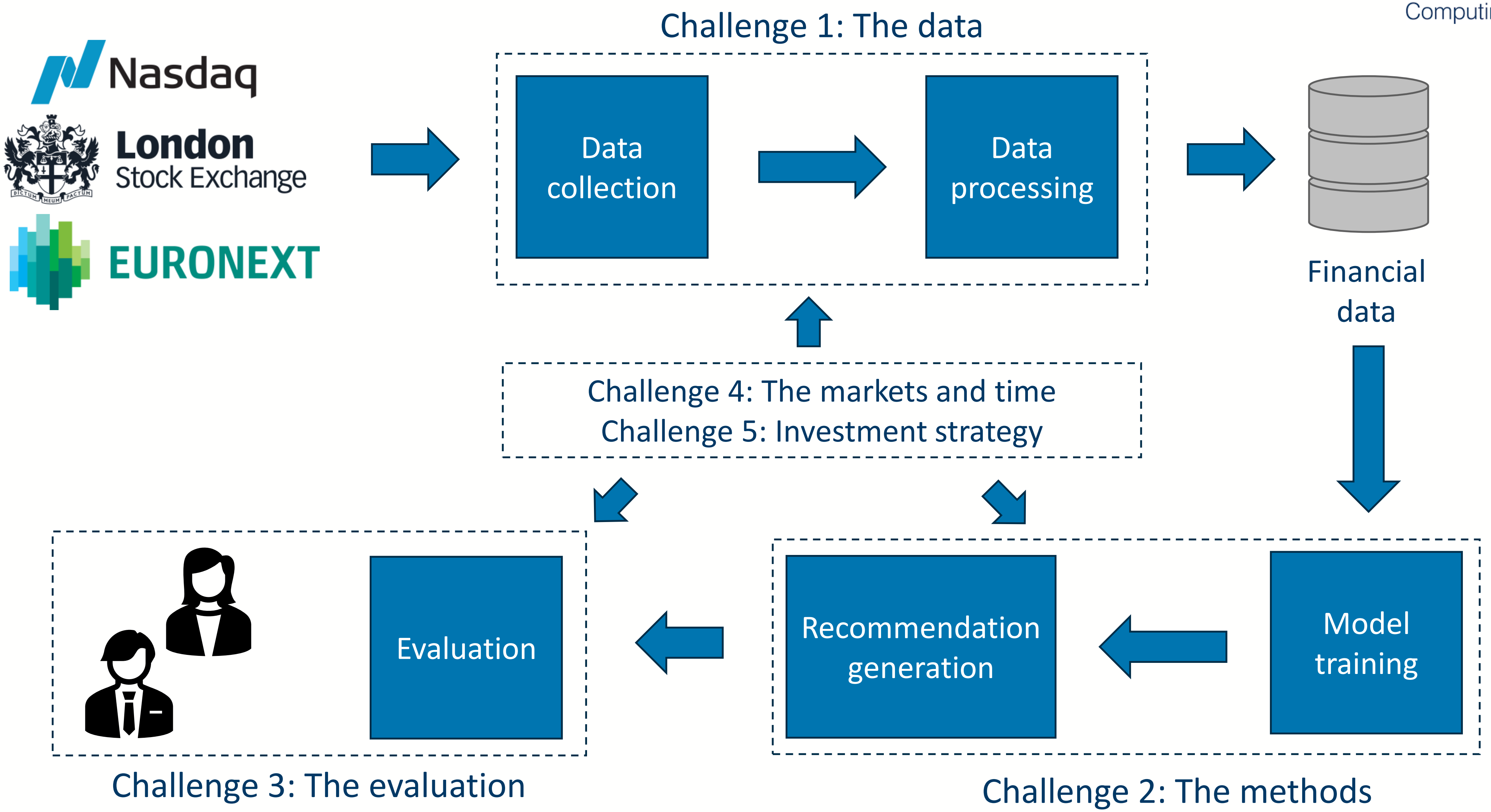
Dr. Javier Sanz-Cruzado

WORLD CHANGING GLASGOW

A WORLD TOP 100 UNIVERSITY

A pipeline for financial asset recommendations

University of Glasgow
School of Computing Science

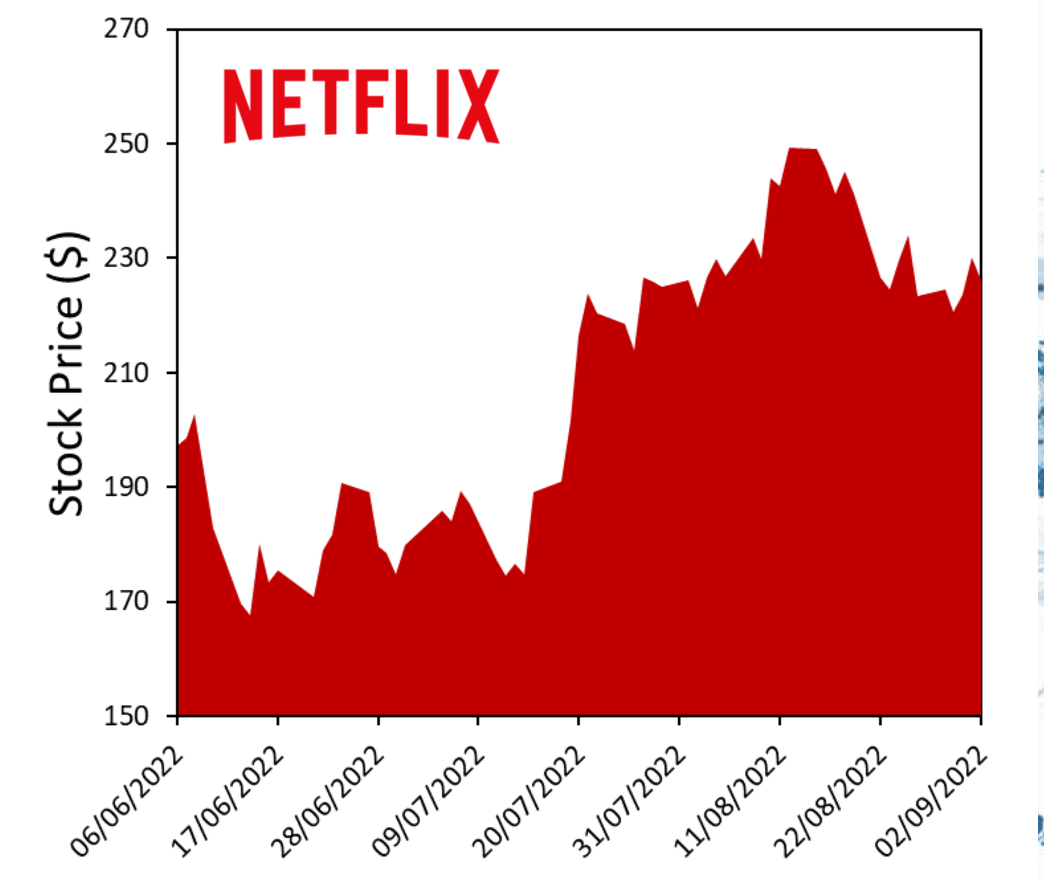


Data

Financial data

University of Glasgow
School of Computing Science

- Fundamental to train, deploy and evaluate financial asset recommendation algorithms
- Properties of financial data:
  - Dynamic
  - Noisy
  - Incomplete
  - Massive
  - Challenging to get



How to get and prepare your data?

University of Glasgow
School of Computing Science

Step 1: Choose financial securities

Exchanges / markets	Asset types

Step 2: Select data types and collect the data from available sources

Step 3: Data cleaning and processing

Step 1a: Financial markets / exchanges

University of Glasgow
School of Computing Science

- Public financial instruments are traded in securities exchanges.
- Every market covers a range of financial products / product types.
- Examples of exchanges:

Nasdaq
NYSE
London Stock Exchange
EURONEXT
BÖRSE BERLIN
BOLSA DE MADRID
SSE SHANGHAI STOCK EXCHANGE
JSE
NSE

- More: https://globalexchangesdirectory.com/

Step 1b: Asset types

University of Glasgow
School of Computing Science

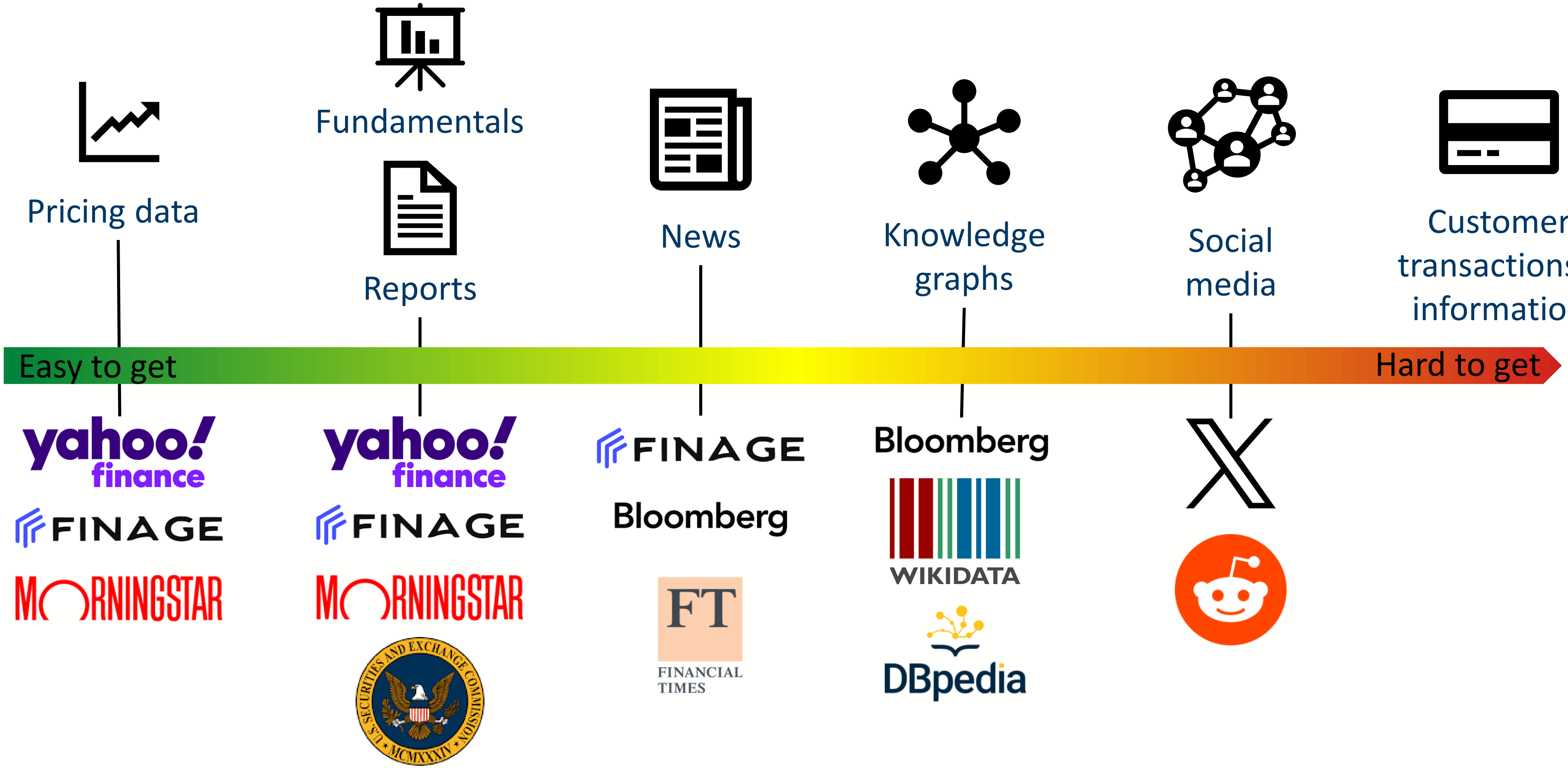
- There are multiple types of securities we can trade in financial markets
- Every security behaves differently

Stocks / Shares	Bonds	Currency (FX)	Exchange Traded Funds	… many more
• Fractional ownership in a company, which usually comes with some voting rights and potentially dividends (pay-outs when the company does well)	• A contract whereby in exchange for money now, the company or government will pay back that money at a later date and pay interest on the borrowed money	• Fiat Currency… or virtual dog-themed pseudo currency	• Fractional ownership of a large pool of other financial assets managed by a company	• Options • Derivatives • Commodities • Fine-art • NFTs

Step 2: Information types

University of Glasgow
School of Computing Science

- We can collect multiple types of information to provide financial asset recommendations



Step 3: Data cleaning

University of Glasgow
School of Computing Science

- Financial data is noisy
- **It is impossible to prevent this!**
- But we can assist our algorithms

- **Warning:** we need to feed our algorithms with good data; otherwise, results might not be as expected!

- **Data cleaning:** check that your data is consistent
  - Don't leak future information!
  - Remove negative / zero prices
  - Make sure closing price is between min and max prices
  - Unify currencies

Step 3: Data processing

University of Glasgow
School of Computing Science

- Our algorithms need to understand the data

- Sometimes, it is necessary to generate features for them to understand it
  - **Technical indicators:** heuristic signals produced by price and volume of a security, used by investors to make decisions
    - Return on investment
    - Volatility
    - Moving average of closing price
  - **News / social media sentiment:** how do people feel towards a particular company / product?
  - **Embeddings:** summarize texts, knowledge graph nodes, etc. in a vector that an algorithm can analyse.

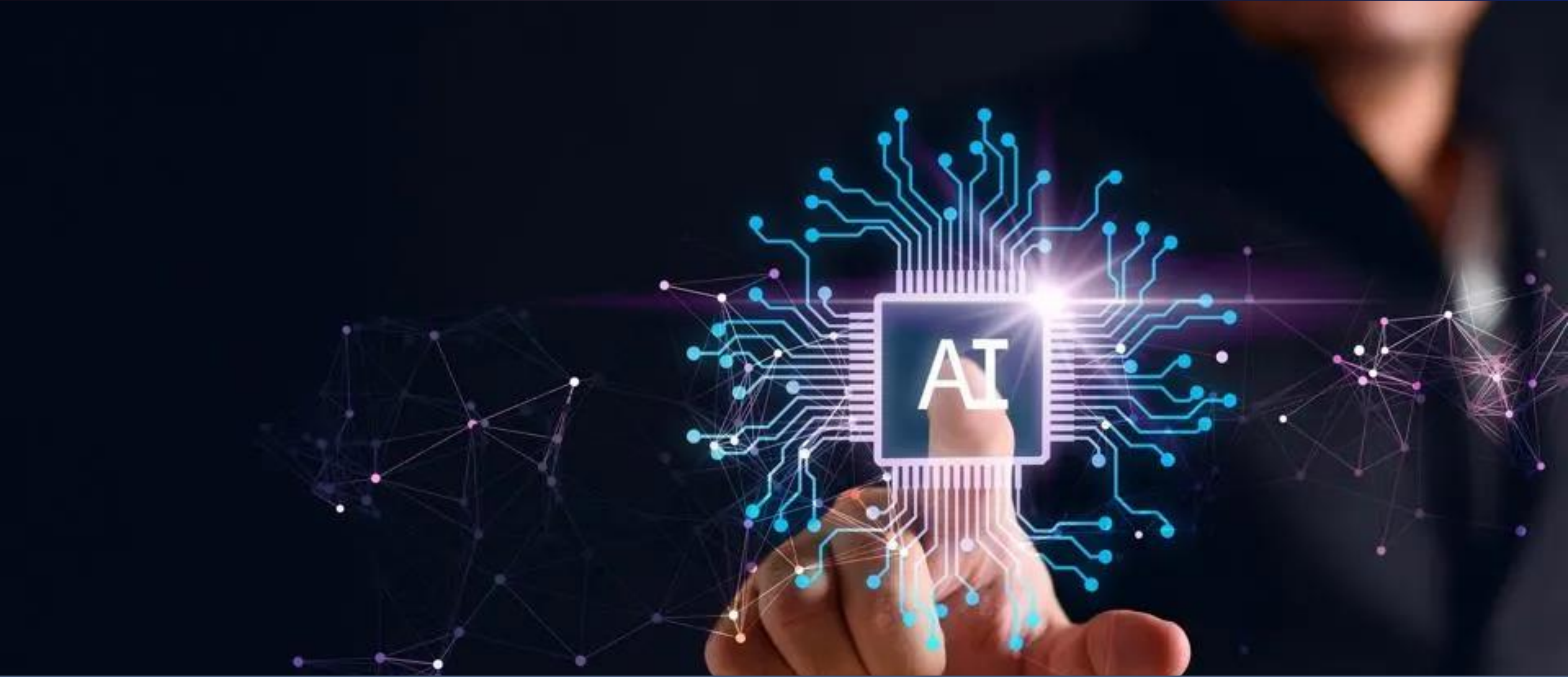
Conclusions on data

University of Glasgow
School of Computing Science

- The data available in the market is massive and noisy

- We need to choose many aspects of it:
  - Securities to consider
  - Types of data

- It is important to clean and process the data
  - Inconsistent data might prevent algorithms from succeeding
  - Algorithms cannot directly use all the information available
  - We need to generate features they can understand



Algorithms

Conclusions on data

University of Glasgow
School of Computing Science

Financial Asset Recommendation (FAR)

Profitability based

- Predict future price / profitability
- Non-personalized

Transaction based

- Uses what others have bought / sold
- Customer / asset similarity

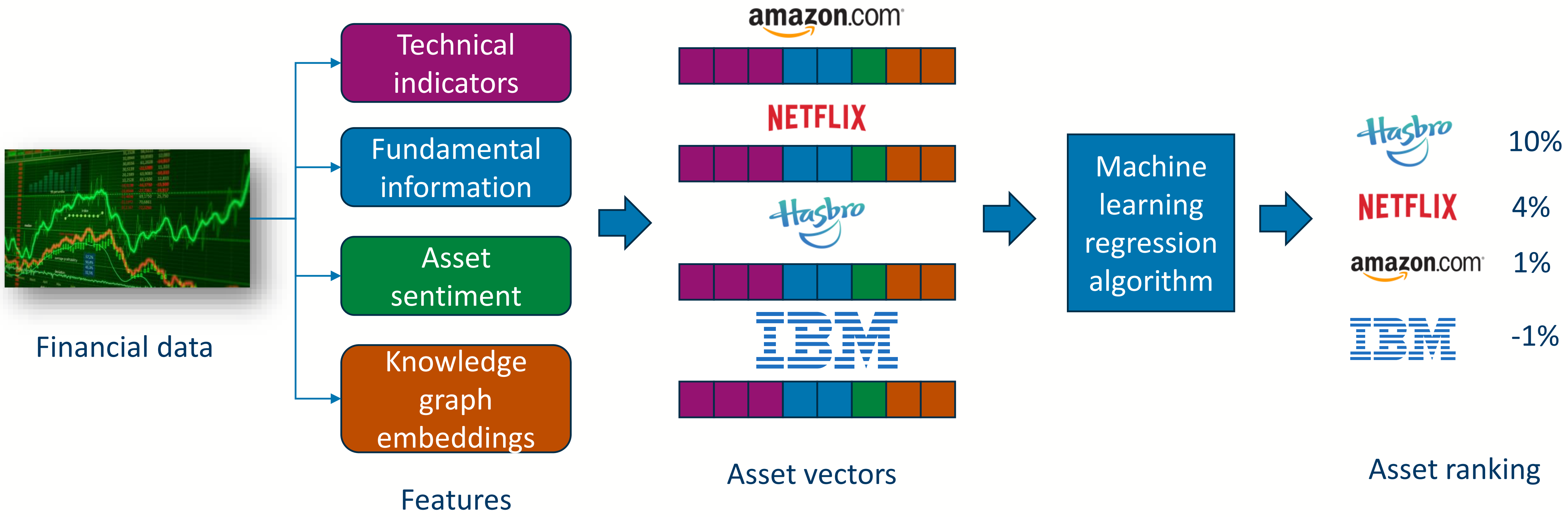
Hybrid

- Combination of previous methods

Profitability-based recommendation

University of Glasgow
School of Computing Science

- **Goal:** predict future profitability / price of financial assets.
- They are based on machine learning regression algorithms.
- **Recommendation score:** estimated profitability at fixed horizon



Profitability-based recommendation

- **Example algorithms:**
  - Linear regression
  - Random forest regression
  - Neural networks (GRU, LSTM)

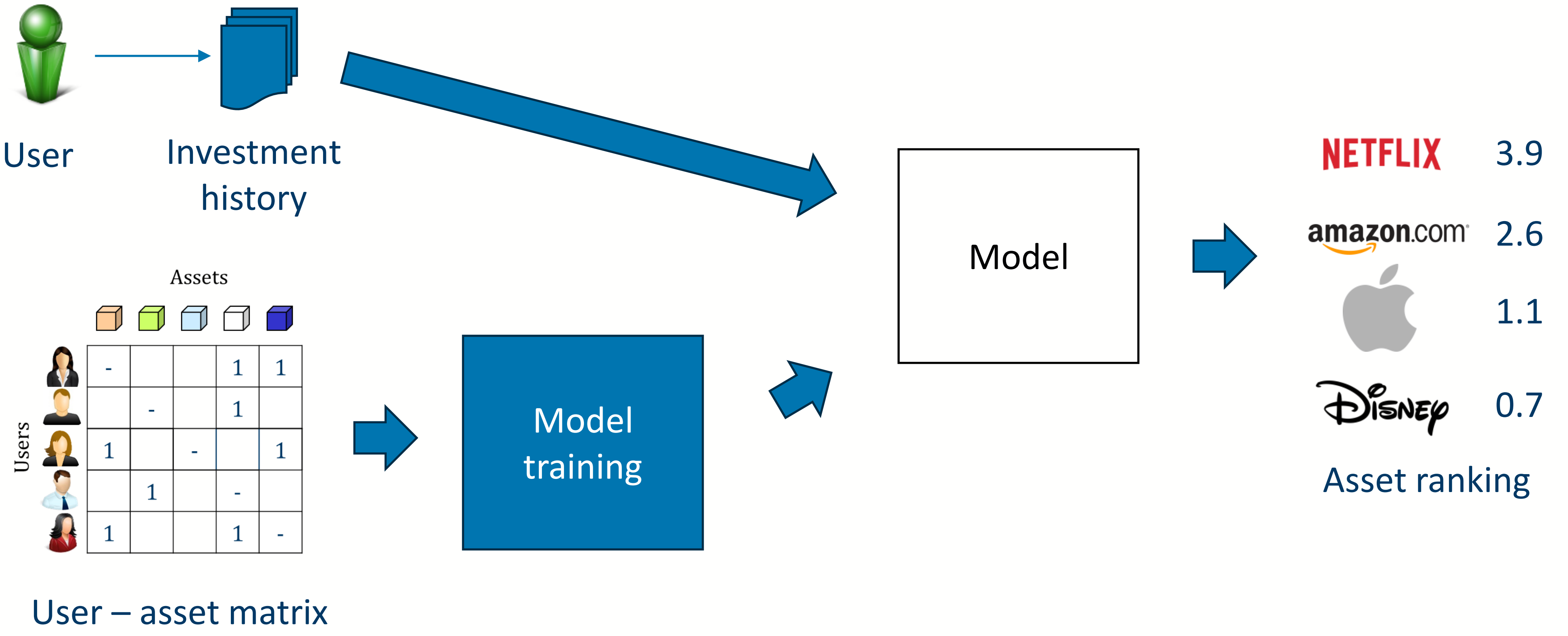
- **Strengths:**
  - Aligned with investor goal of earning money.
  - Capable of dealing with time series information.
  - Capable of integrating multiple sources of information in a simple way.

- **Weaknesses:**
  - Not personalized (only use asset information)

Transaction-based recommendation

University of Glasgow
School of Computing Science

- Consider customer interactions with financial assets.
- **Goal:** Predict future investments of the investors.
- **Recommendation score:** estimated utility of an asset for an investor.



Types of algorithms

University of Glasgow
School of Computing Science

- Collaborative filtering
  - **Idea:** similar customers invest on similar assets.
  - **Algorithms:** k-nearest neighbors, matrix factorization, LightGCN

- Content-based:
  - **Idea:** customers invest on similar assets to those they invested in the past
  - Create a customer profile representation from asset information about past investments.

- Demographic recommendation:
  - **Idea:** customers with similar demographic features (age, risk aversion, etc.) invest on similar assets
  - **Algorithms:** k-nearest neighbors

Types of algorithms

University of Glasgow
School of Computing Science

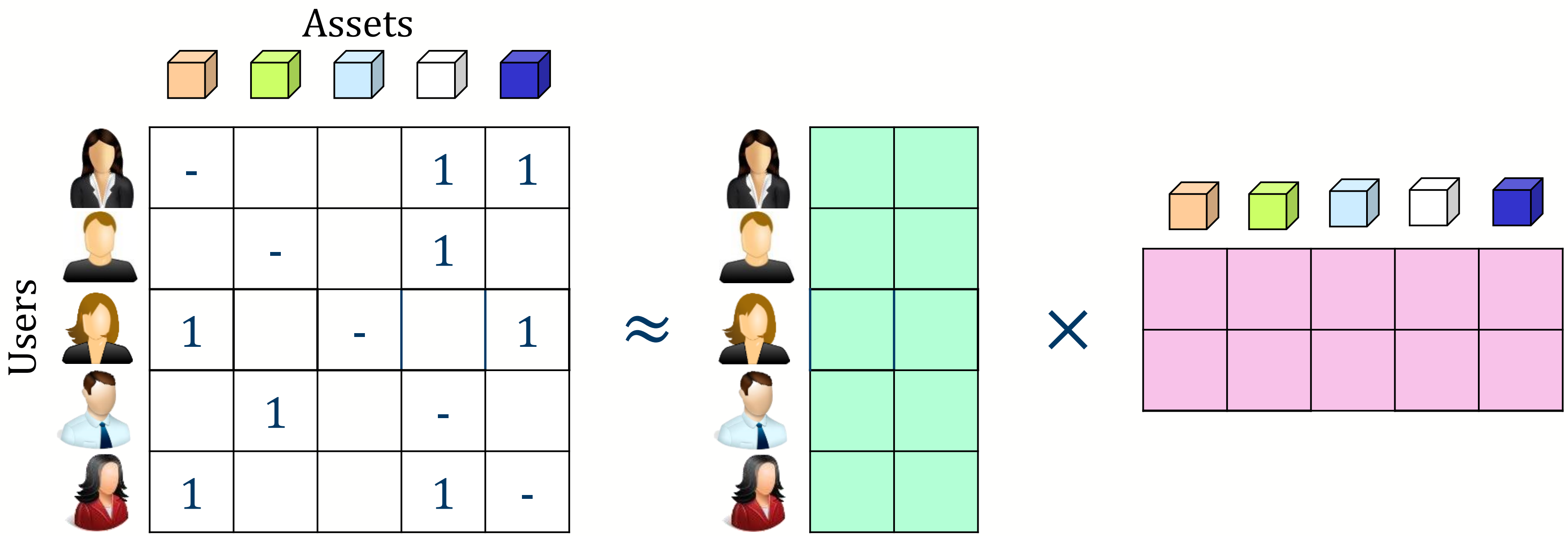
- Social-based recommendation
    - **Idea:** social connections influence users on their investments.
    - **Algorithms:** trust-aware k-nearest neighbors.

- Knowledge-base recommendation:
    - **Idea:** apply specific domain knowledge about how items meets user needs and preferences.
    - **Algorithms:** case-based reasoning, fuzzy logic.

Algorithm example: matrix factorization

University of Glasgow
School of Computing Science

- Matrix factorization represents users and items in a low-dimensional latent space.



- User vectors summarize customer preferences on assets.
- Item vectors do the same with assets.
- **Scores:** product on customer / asset vectors.

Transaction-based recommendations

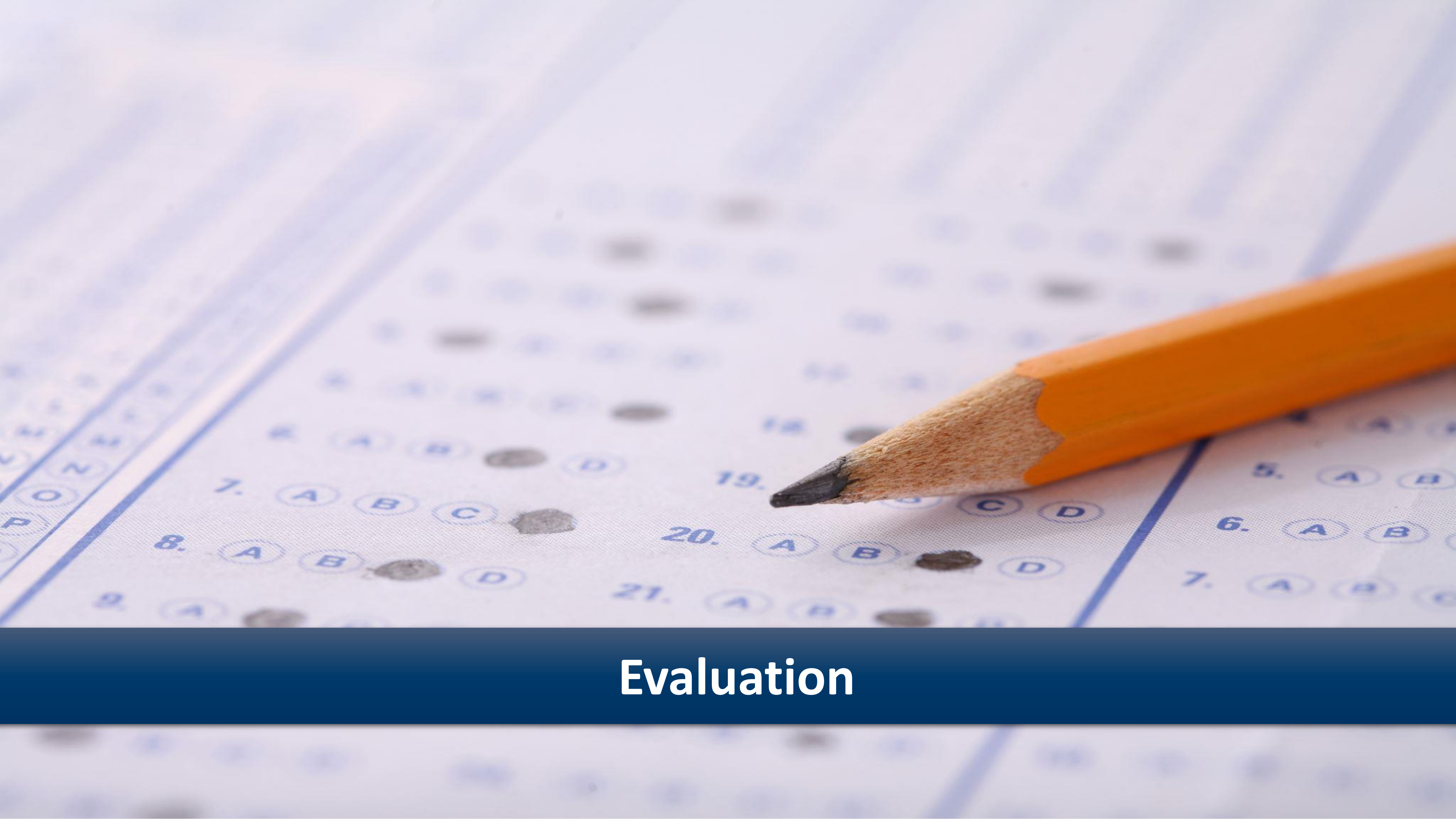
- **Strengths:**
  - Personalized (deal with customer information).
  - They capture customer interests.

- **Weaknesses:**
  - Don’t consider pricing information.

Hybrid-based recommendations

University of Glasgow
School of Computing Science

- Combine different types of data
- For instance, we could build a model using recommendation scores from other algorithms
- **Idea:** have all the strengths of previous algorithms without any of the weaknesses



Evaluation

Evaluation Goals

University of Glasgow
School of Computing Science

01

Determine whether recommendations are **useful**

02

Identify 'blind spots' in solutions

03

Choose the **best approaches** for deployment

How do we evaluate?

- Essentially, depends on the goals of our system
- … or the goals of our customer

- An investor wants to increase their money by using the system.
- A financial institution wants to increase their revenue by having more customers investing on the system.

How do we evaluate?

University of Glasgow
School of Computing Science

Profitability based ← **Financial Asset Recommendation** → **Transaction based**

Do our customers earn money?

- Aligned with customer interests
- Ignores past/future customer actual investments

- **Metrics:** Key performance indicators at a fixed time interval
  - Return on investment (ROI)
  - Net profit

Can we predict future investments?

- Investment transactions indicate strong preference
- Relevant transactions: acquisitions
- Ignores temporal pricing information

- **Metrics:** Recommender systems metrics
  - Precision
  - nDCG

Ideal Scenario

Recommendation algorithm

recommends

Profitable assets

AND

Assets customers like

Profitability-based metric example

University of Glasgow
School of Computing Science

Return on investment (ROI)

- **For an asset:** relative variation of the price of the asset between the recommendation time t and a fixed period Δt afterwards

$$\mathrm{ROI}(i, t, \Delta \mathrm{t}) = \frac{\mathrm{price}\ (i, t + \Delta t)\ - \mathrm{price}(i, t)}{\mathrm{price}(i, t)}$$

- **For a recommendation:** average variation of price on the recommended assets.

- Equivalent to profitability of a fund where we invested equally on every asset.

Profitability-based metric example

University of Glasgow
School of Computing Science

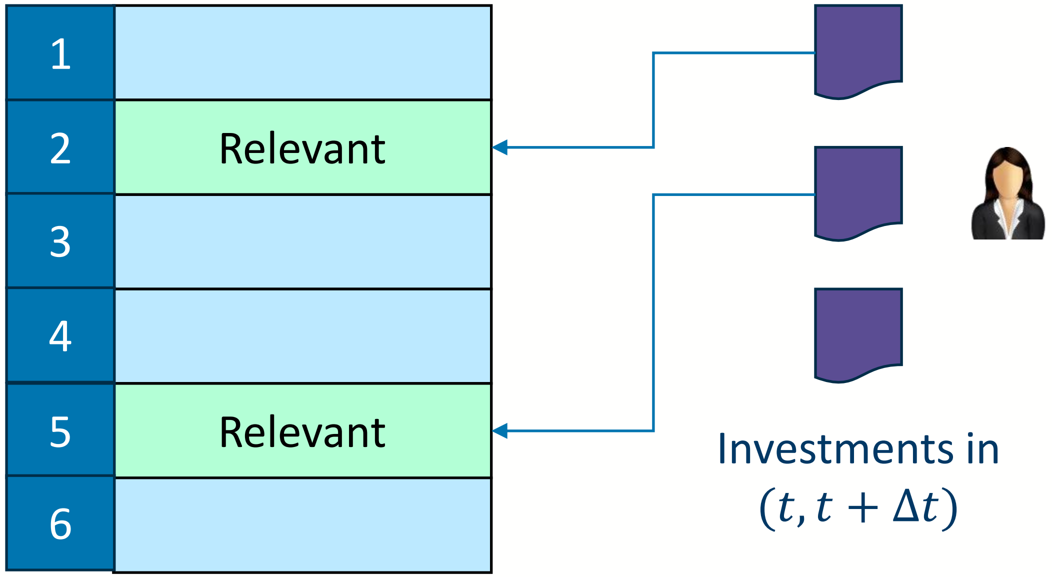
Monthly Return on investment (ROI)

- **Problem:** ROI is impossible to compare among different time horizons.
- **Solution:** normalize it – estimate, for instance, how much price moves every month.

$$\mathrm{ROI}(i, t, \Delta\mathrm{t}) = \left(1.0 + \frac{\mathrm{price}\ (i, t + \Delta t)\ - \mathrm{price}(i, t)}{\mathrm{price}(i, t)}\right)^{\frac{30\ \mathrm{days}}{\Delta t}} - 1.0$$

Transaction-based metric example

Normalized cumulative discounted gain (nDCG)



Recommendation ranking R

An item i is relevant for a customer u if u has invested in i in the $(t, t + \Delta t)$ period

$$\mathrm{DCG@}k(u, R) = \sum_{i=1}^{k} \frac{\mathrm{rel}(u, i, t, \Delta t)}{\log_2(1 + i)}$$

We prefer relevant assets in the first ranking positions

Normalize between the maximum possible value

$$\mathrm{IDCG@}k(u) = \max_{R}\left(DCG@k(u, R)\right)$$

$$\mathrm{nDCG@}k\ (u, R) = \frac{DCG@k(u, R)}{IDCG@k(u)}$$

How good are recommendation algorithms?

- We conduct experiments over **real financial data**

- We **test different financial asset recommendations algorithms**
  - **Price-based** algorithms
  - **Transaction-based** algorithms

- Are they profitable?
- Do they capture customer preferences?

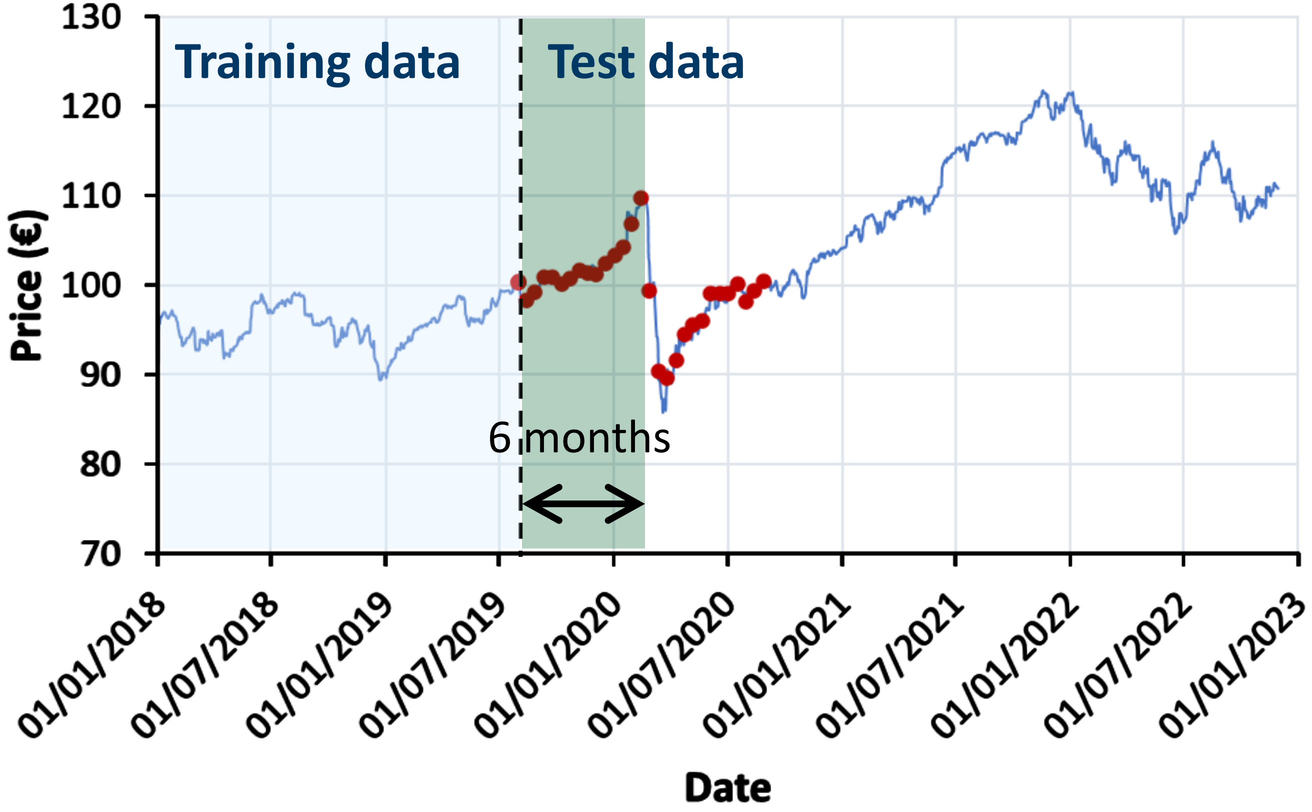
Dataset

University of Glasgow
School of Computing Science

- **Greek market:** stocks, bonds, mutual funds
- **Period:** 1st January 2018 – 30th November 2022
- **Combines**
  - Time series data (pricing information)
  - Customer investments
- **Time series data:**
  - 807 financial assets (321 assets with investments)
  - 703,303 data points
- **Customer investments:**
  - 29,091 customers
  - 387,783 transactions

Experimental procedure

University of Glasgow
School of Computing Science



Procedure

1. Select recommendation time t
2. Split into training / test
   - **Training:** from 1st Jan 2018 to t
   - **Test:** from t to $t + 6$ months
3. Train models
4. Execute recommendations at t
5. Evaluate

29 time splits

- One every two weeks
- From: 1st August 2019
- To: 28th August 2020

Profitability Prediction Algorithms

University of Glasgow
School of Computing Science

- Only based on price time series

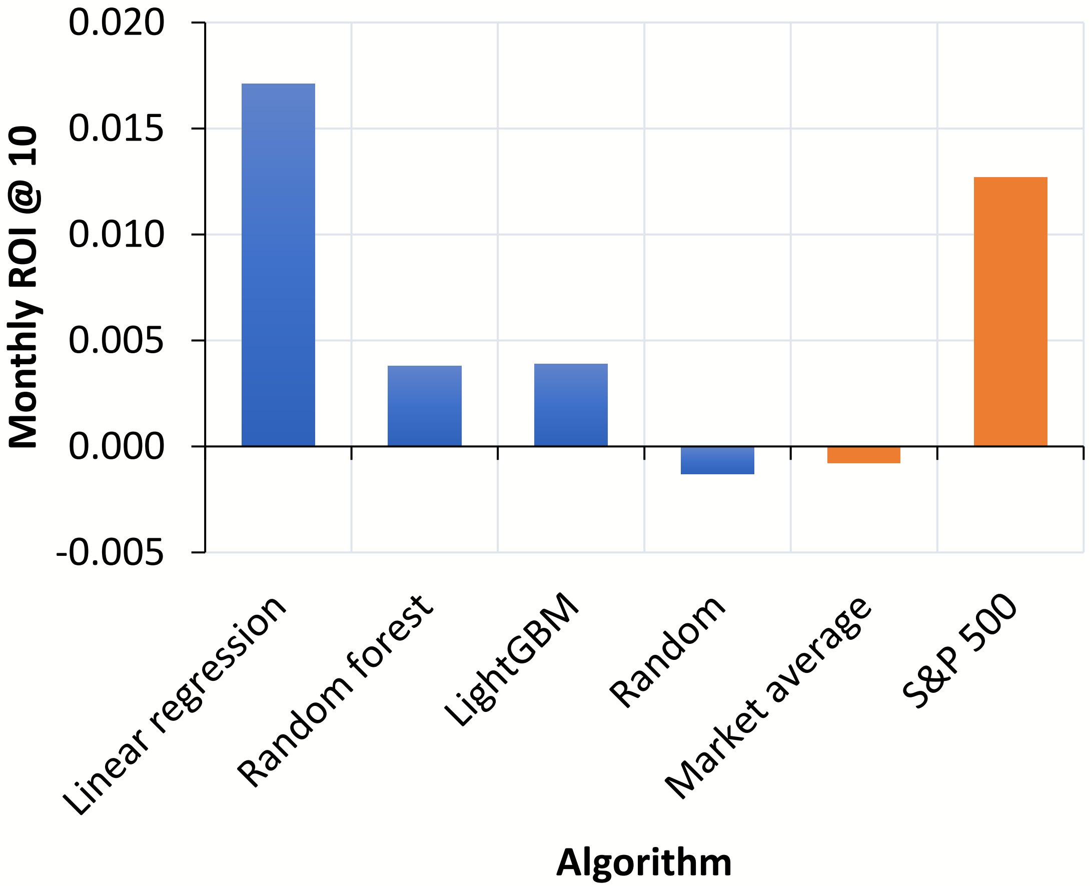
- Use financial technical indicators to predict future profitability
  - ROI at 1, 3, 6 months
  - Volatility at 1, 3, 6 months
  - Etc.

- Three methods:
  - Linear regression
  - Random forest
  - LightGBM

How good are profitability prediction algorithms?

University of Glasgow
School of Computing Science

We compare the algorithms in terms of monthly ROI over the **top 10** recommended results



- All three algorithms improve the market average and random recommendation
- Linear regression improves S&P 500 index

Transaction-based models

University of Glasgow
School of Computing Science

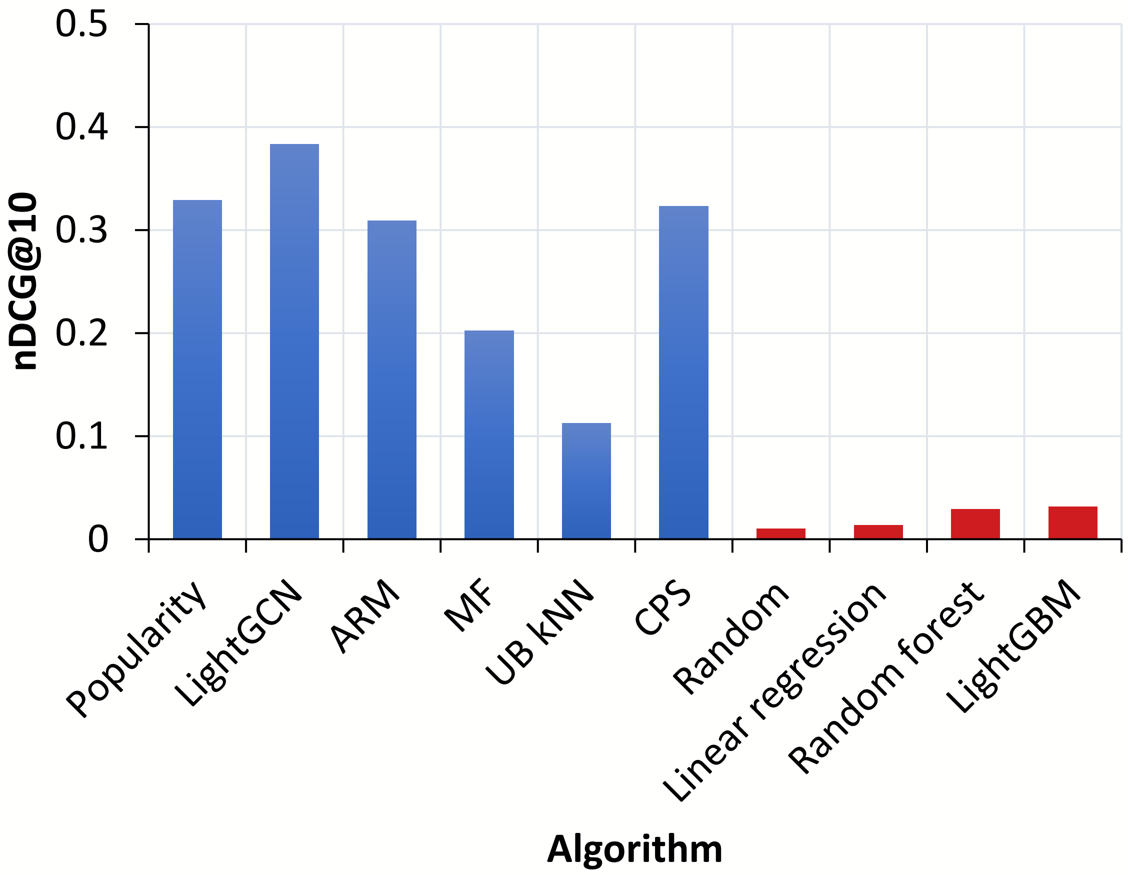
- Based on **past transactions of the customer**
- Consider what other customers have invested in

- Methods
  - Non-personalized: **Popularity**
  - **Collaborative Filtering**:
    - LightGCN
    - ARM
    - User-based kNN
  - **Demographic Information**: CPS

How good are Transaction-based Algorithms Predicting Customer Tastes?

University of Glasgow
School of Computing Science

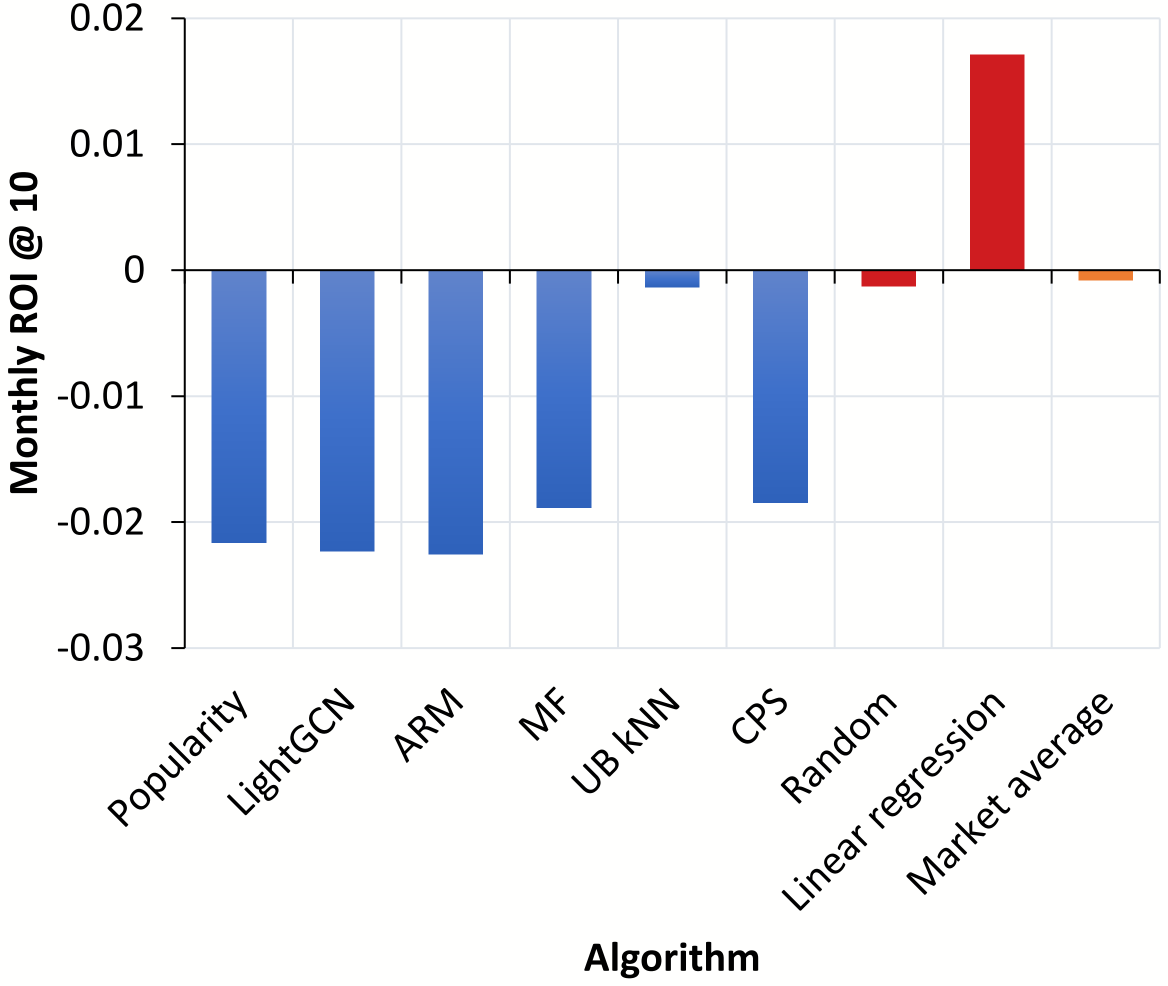
We look, again, at the top 10 recommended results



- All transaction-based algorithms recommend assets that are much more similar to what the customers actually later invested in
- But, are these recommendations profitable?

How good are Transaction-based Algorithms Recommending Profitable Assets?

University of Glasgow
School of Computing Science



- Transaction-based algorithms **suggest (in general) unprofitable assets**
- They don't beat the market (on average)
- Profitability prediction algorithms are better at recommending profitable assets
- But transaction-based models capture better the behaviour of customers
- Which algorithms are better?

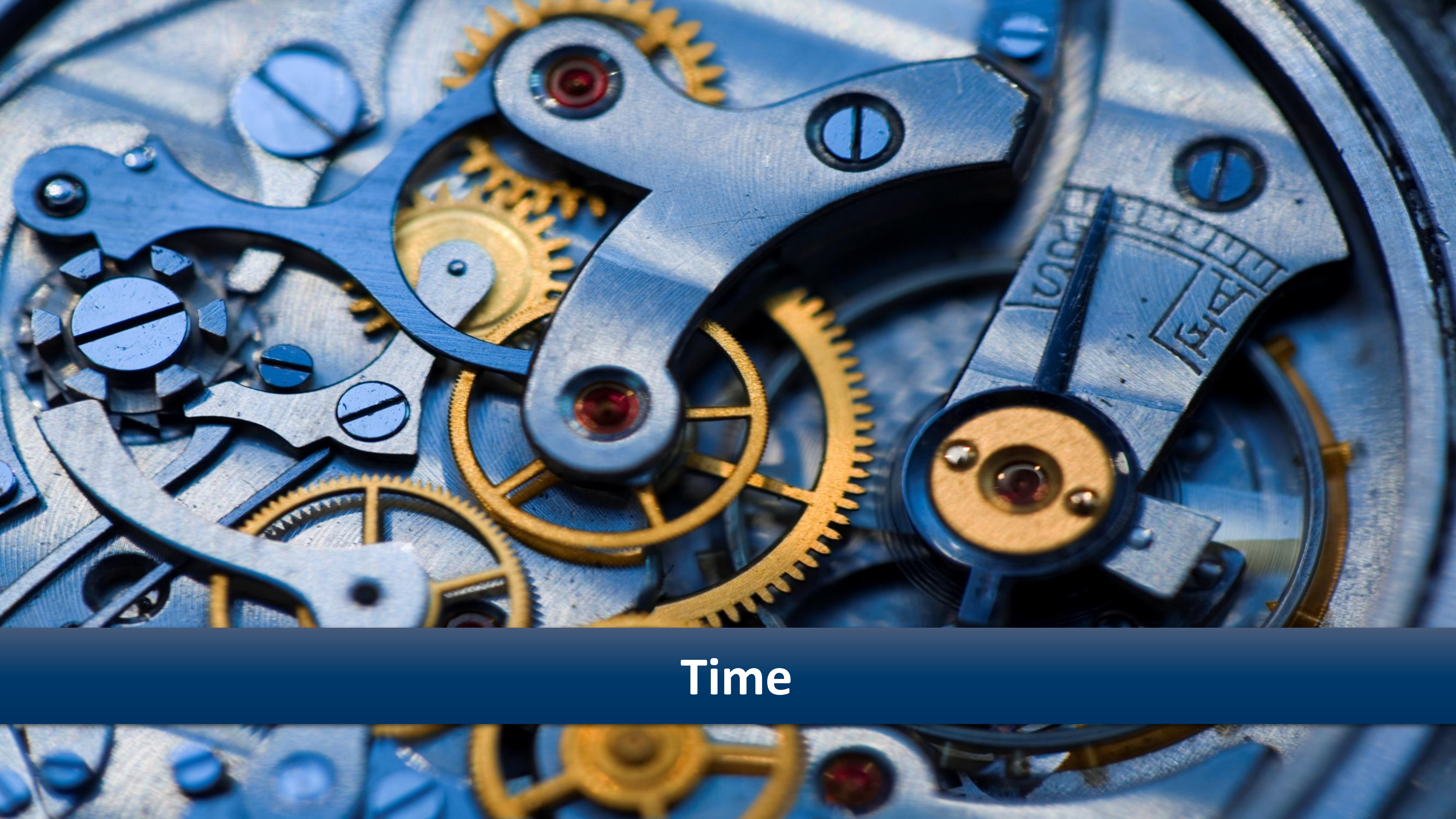
Conclusions on evaluation

University of Glasgow
School of Computing Science

- We need evaluation to determine how good our systems are.

- There are two main ways to evaluate financial asset recommendations
  - Profitability: do our customers earn money?
  - Relevance: are we able to predict future investments?

- We might choose different algorithms depending on our goal:
  - Profitability prediction algorithms work better for finding profitable assets
  - Transaction-based methods are better at identifying future customer investments
  - But the opposite is not necessarily true



Time

Time

University of Glasgow
School of Computing Science

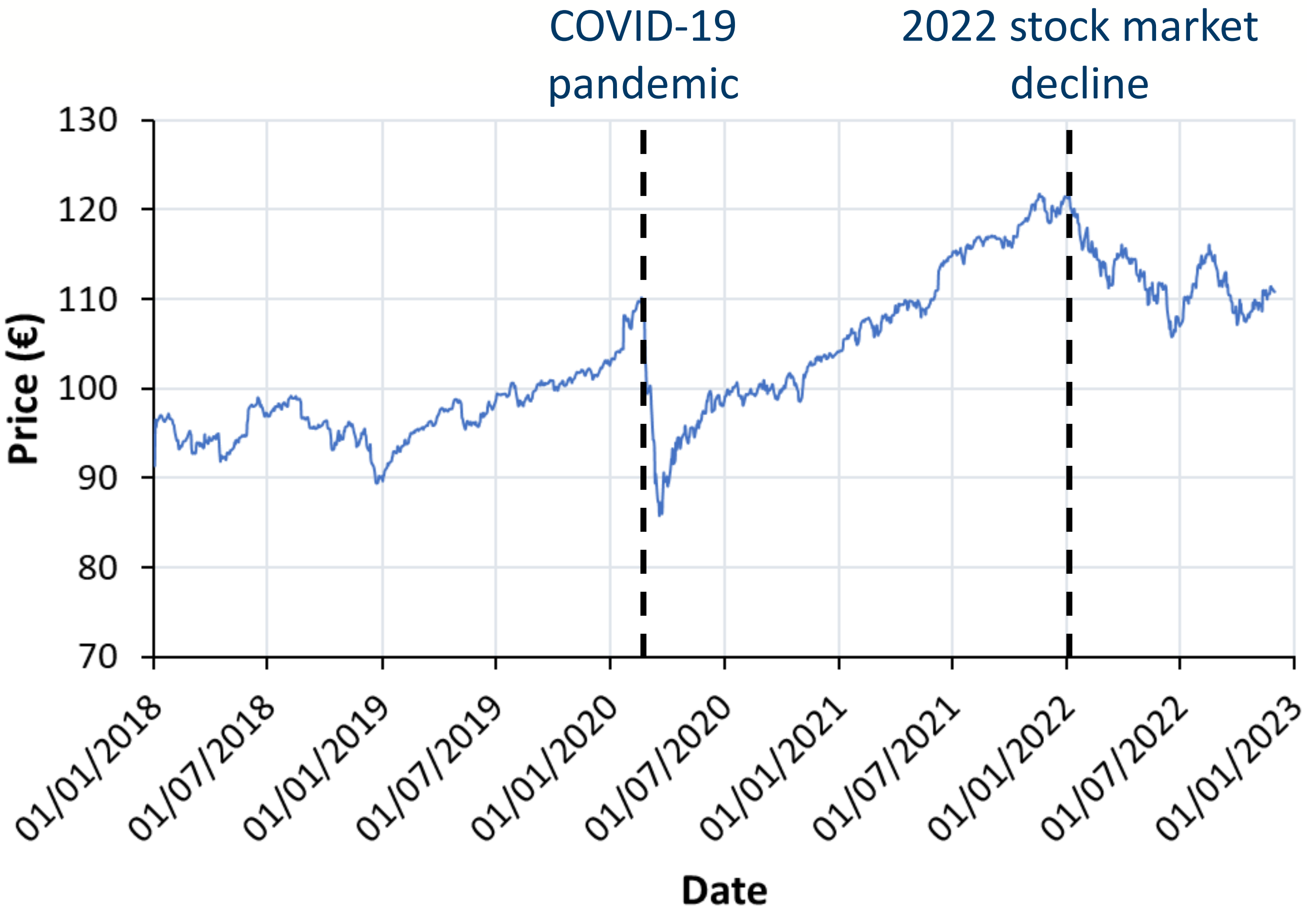
- In our previous experiments, we averaged our results over 28 different dates

- But, why? Isn’t enough to check on one date?

- The answer is no:
  - Financial markets are very dynamic
  - Asset prices might change every few seconds (or less!)
  - Markets are also affected by external events
    - Product releases
    - Global events: pandemics, wars can affect the whole market!

An example on the Greek market

University of Glasgow
School of Computing Science

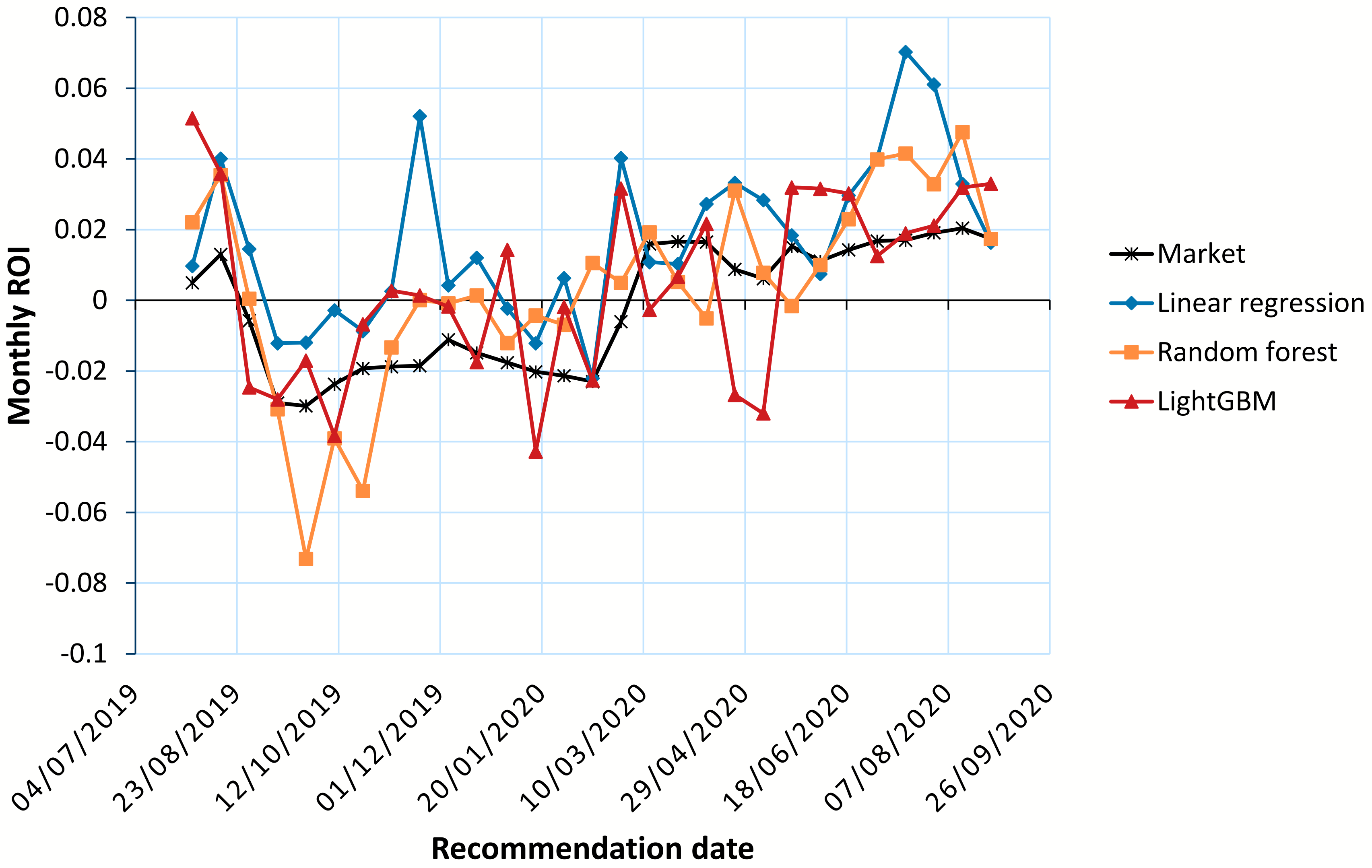


- Prices are not stable
- Market grows and declines depending on multiple factors
- Clear examples:
  - Covid 19 pandemic
  - 2022 stock market declined (aggravated by Ukraine war)
- Does this affect algorithms?

Profitability over time (profitability-based methods)

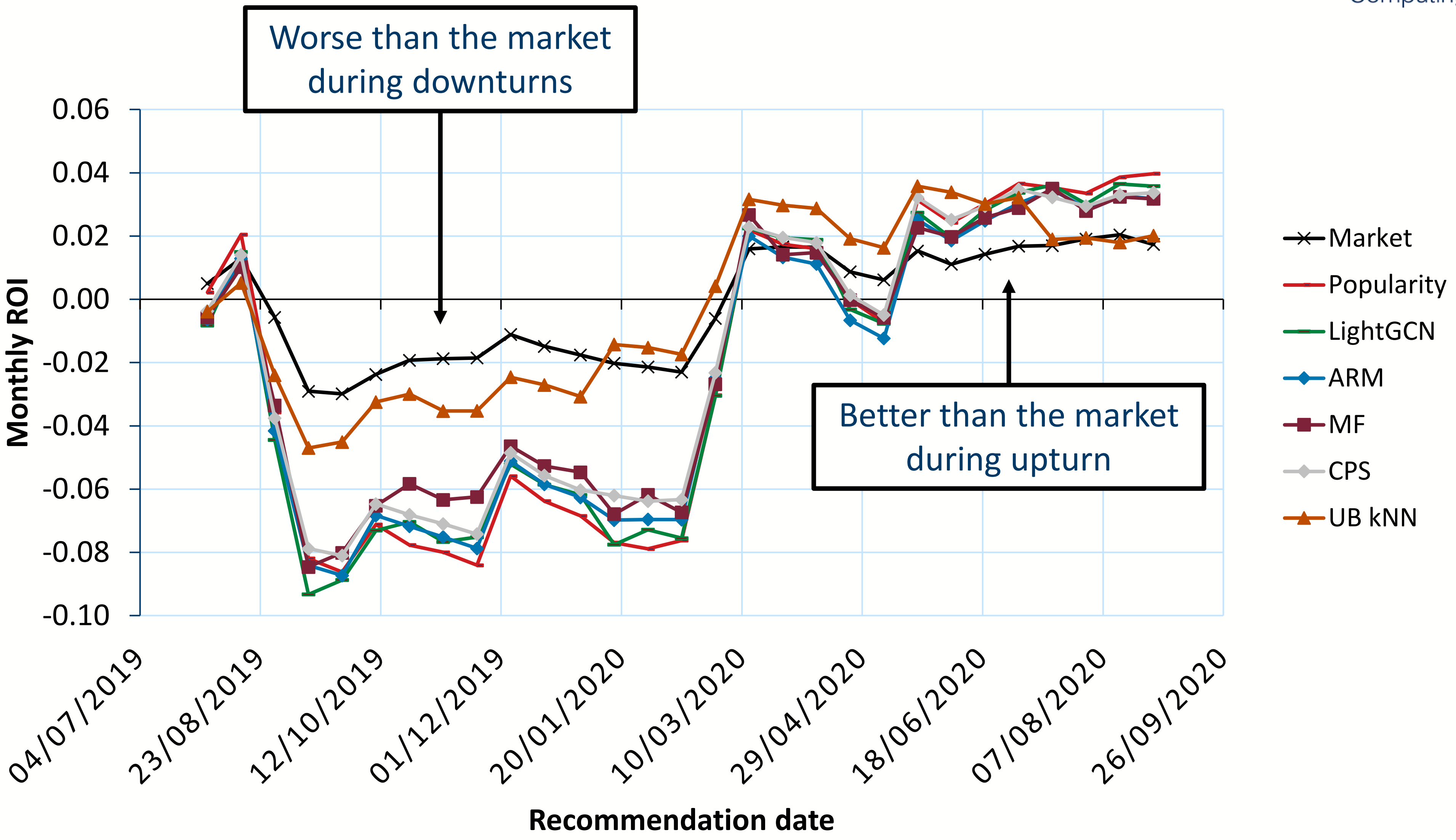
University of Glasgow
School of Computing Science

We look 6 months into the future from this dates!



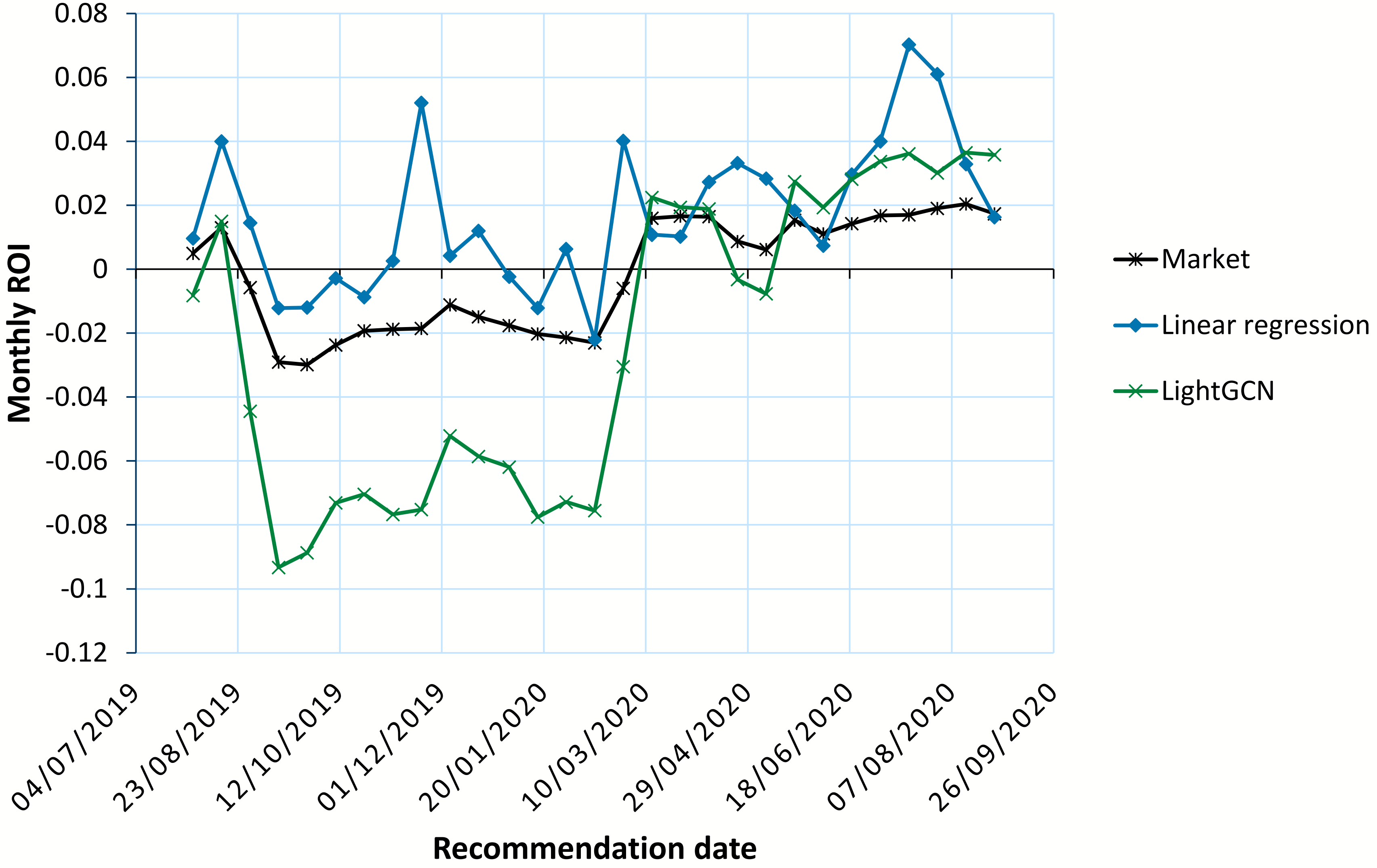
Profitability Over Time (transaction-based models)

University of Glasgow
School of Computing Science



Profitability over time (comparison)

University of Glasgow
School of Computing Science



Conclusions

University of Glasgow
School of Computing Science

- In financial asset recommendations, time matters.
- Markets change a lot over time.

- Different financial asset recommendations might work very differently depending on the date we test it.
- If we test it over a single date, we might risk not detecting unwanted behaviours.

- **Solution:** consider different market conditions when training / evaluating
  - Upturns
  - Downturns
  - Stable periods

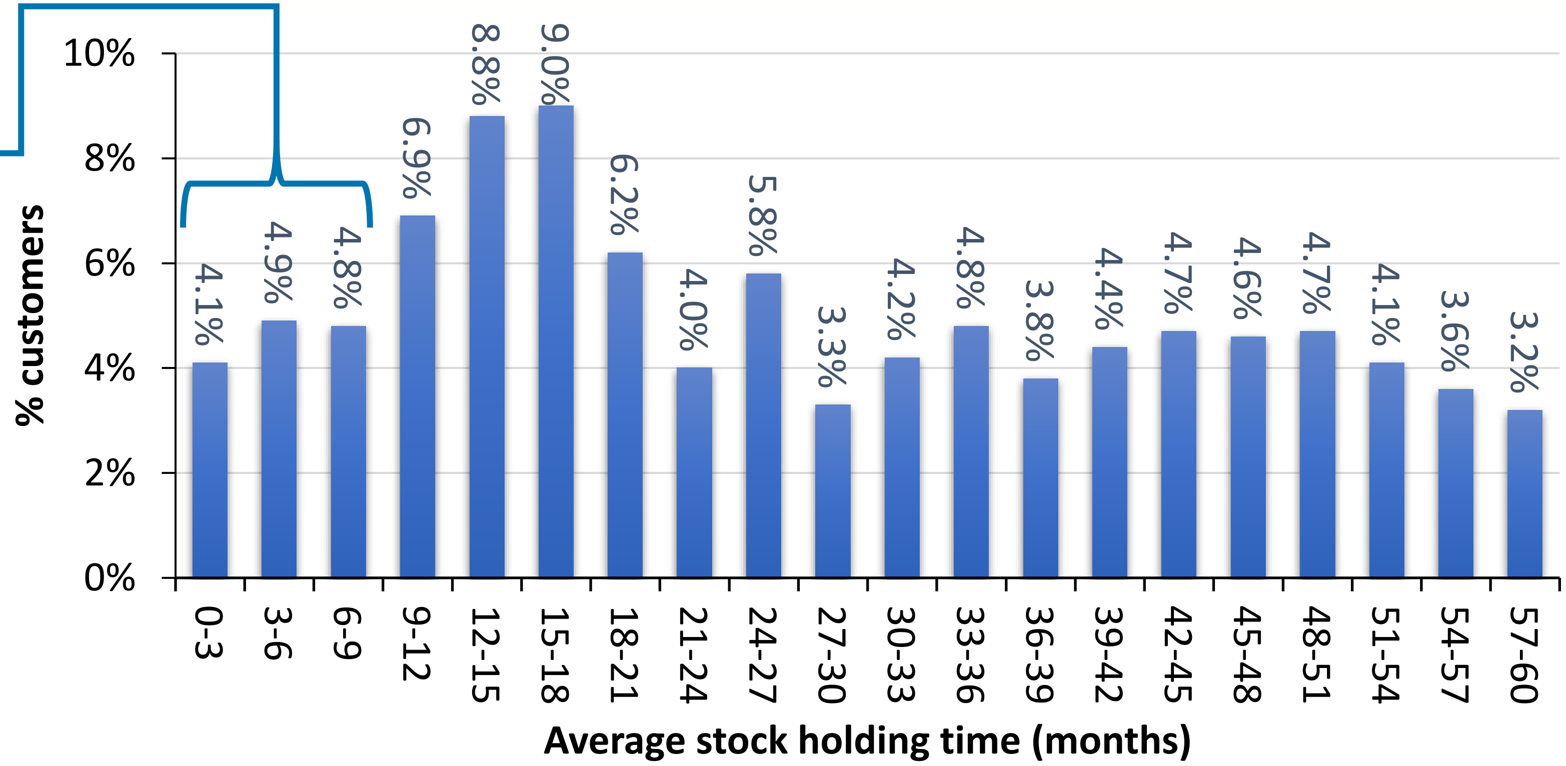
Investor preferences

Investors

University of Glasgow
School of Computing Science

- Investors might follow different strategies on their assets
- How long do investors hold their assets?
- In our previous experiments, we assumed 6 months, but is that realistic?

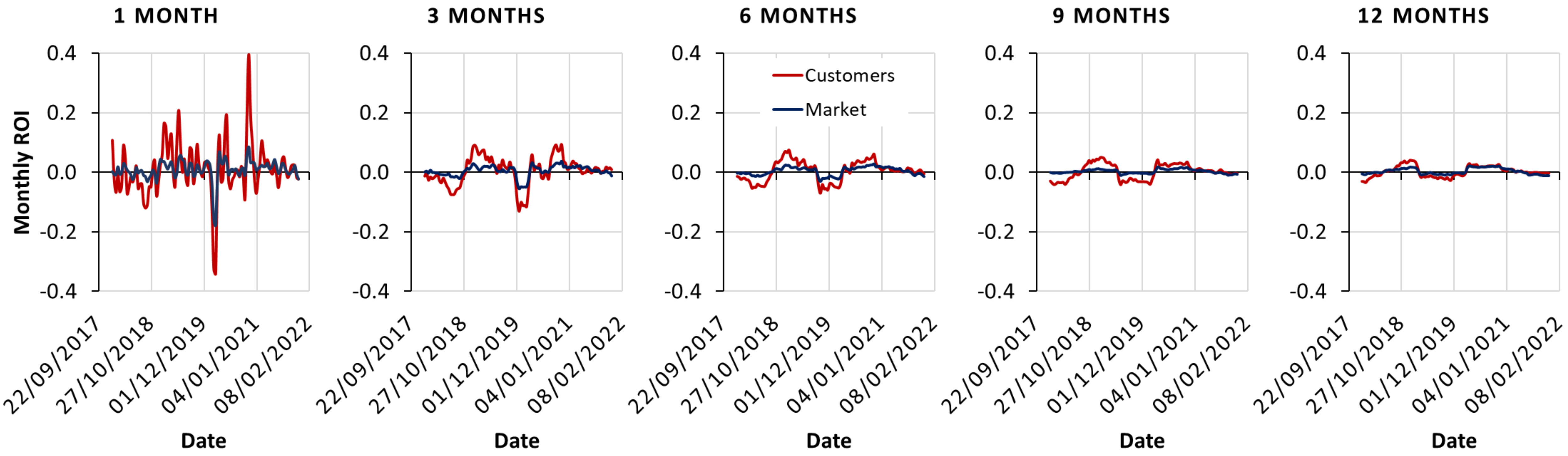
Only around 13% of customers hold their investments for less than 9 months!



Investment horizons in our data might not align with just 6 months

Monthly Profitability at Different Time Horizons

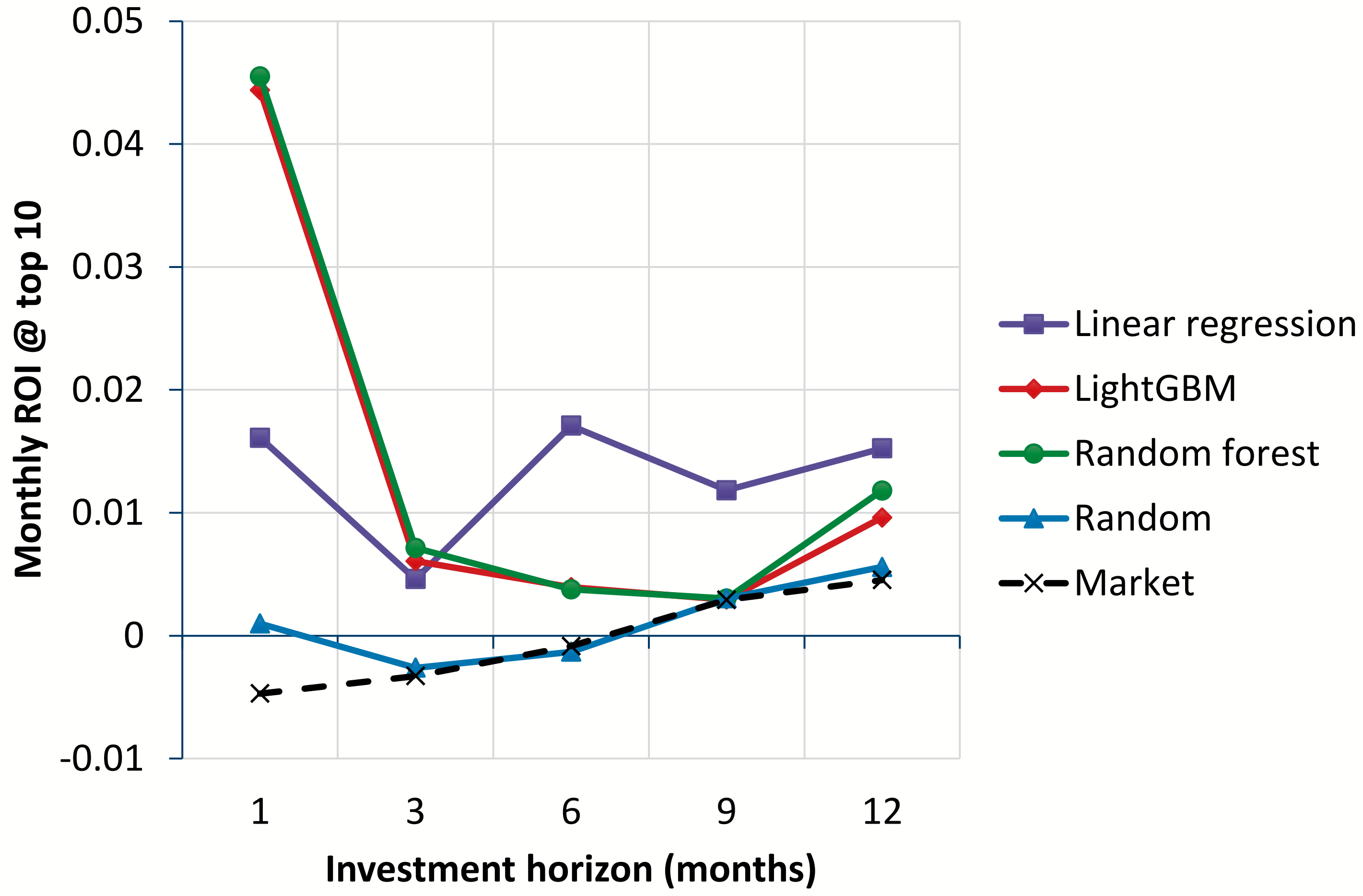
University of Glasgow
School of Computing Science



- The further we look into the future, the smaller (in absolute value) changes are
- What happens if we train models at different time horizons?

Price Prediction for Different Time Horizons

University of Glasgow
School of Computing Science



- Different algorithms work better at different time horizons
- Random forest / LightGBM are more effective at shorter time horizons
- Linear regression is consistent over time horizons (and better at longer term)

Conclusions

University of Glasgow
School of Computing Science

- In financial asset recommendations, we need to consider investor preferences.
- We might need to train different models for different preferences.

- We have seen this, for example, with the investment horizon.
- But there might be other factors:
  - Risk aversion
  - ESG preferences
  - Etc.

Conclusions

Conclusions

University of Glasgow
School of Computing Science

- We have explored five challenges of financial asset recommendations
  - Gather data
  - Choose a model
  - Evaluate the recommendations
  - Effect of time
  - Effect of investor strategies

- We have analysed the effectiveness of two groups of algorithms
  - Pricing-based methods represent promising algorithms, as they help customers beat the market
- Transaction-based methods capture customer preferences…
  - … but recommend non-profitable assets, making them unreliable

- Best methods might change depending on the time / customer investment horizon

Automatic Recommendation… Recommendations

- Clean your data
  - Financial data is noisy
  - Bad data can hurt performance

- Train and evaluate models on varying market conditions
  - Some models might only work during upturns
  - But lose money during downturns (i.e. COVID-19)

- Consider investment strategies in design and evaluation

Questions?

Dr. Javier Sanz-Cruzado

Financial Recommendation Systems

javier.sanz-cruzadopuig@glasgow.ac.uk