Exploiting Knowledge Graph Embeddings For Profitability Prediction on Financial Asset Recommendations

University of Glasgow | School of Computing Science

Terrier Team

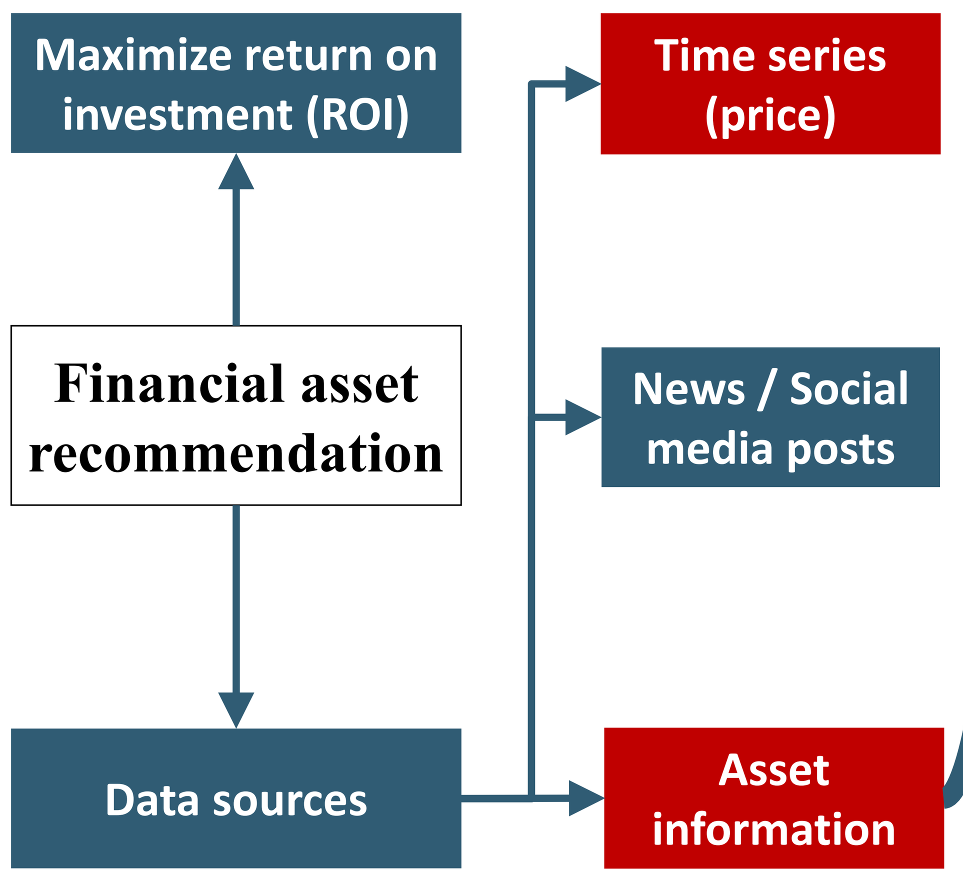
Infinitech
https://www.infinitech-h2020.eu/

Lubingzhi Guo, Javier Sanz-Cruzado, Richard McCreadie, Craig Macdonald, Iadh Ounis
l.guo.1@research.gla.ac.uk,
{javier.sanz-cruzadopuig,richard.mccreadie,craig.macdonald,iadh.ounis}@glasgow.ac.uk

Overview

Financial Asset Recommender (FAR) systems suggest investment assets to customers based on past information. Knowledge graphs represent another source of information, showing (a) information about a particular asset and (b) relations between different assets. In this work, we explore the importance of knowledge graphs for improving recommendations by integrating knowledge graph embeddings as features for profitability estimation algorithms.

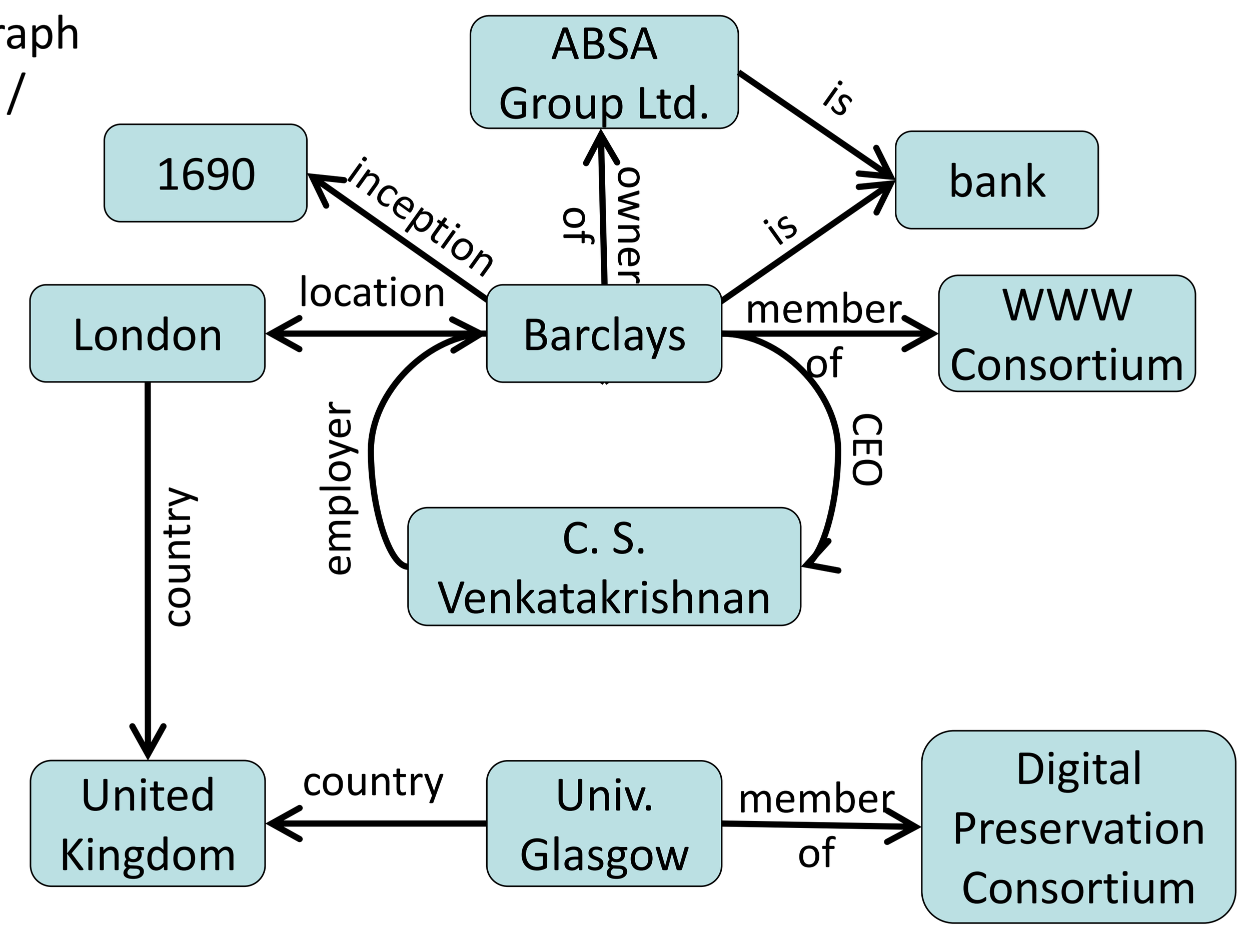
1. Task and Motivation



Knowledge graphs

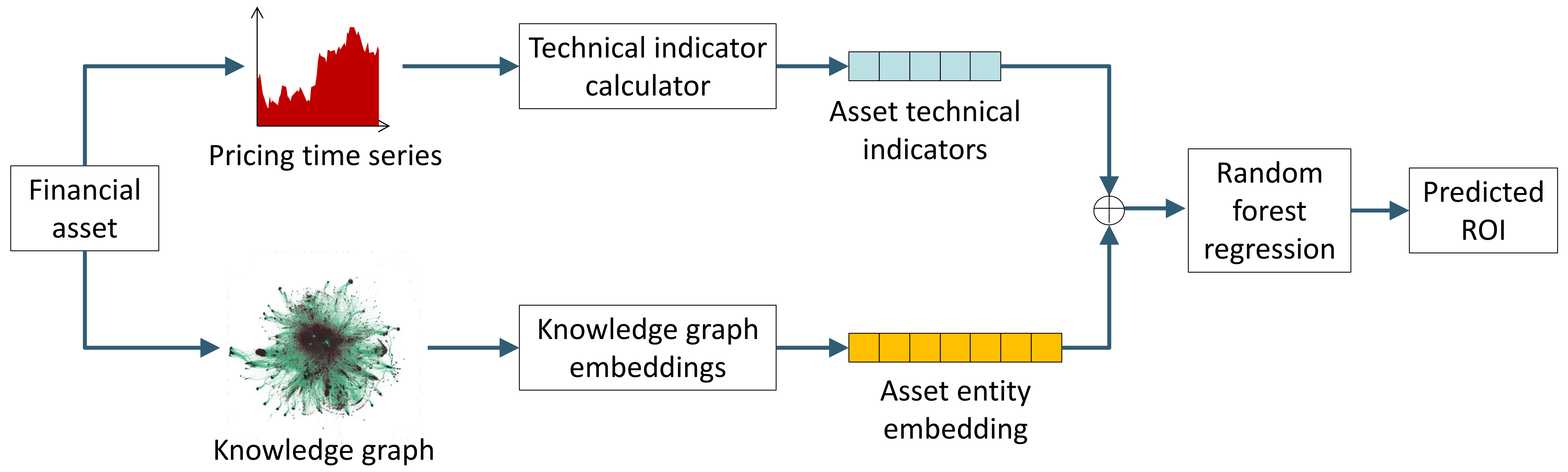
- Representation of the real world as a graph
- Nodes / Entities: represent real objects / concepts.
- Edges: relations between entities
- How can we use them in FAR?
- **Knowledge graph embeddings**
  - Represent nodes as dense vectors
  - Encode both nodes and relations

Knowledge graph example



Research Question: *Can we use knowledge graph embeddings to improve financial asset recommendations?*

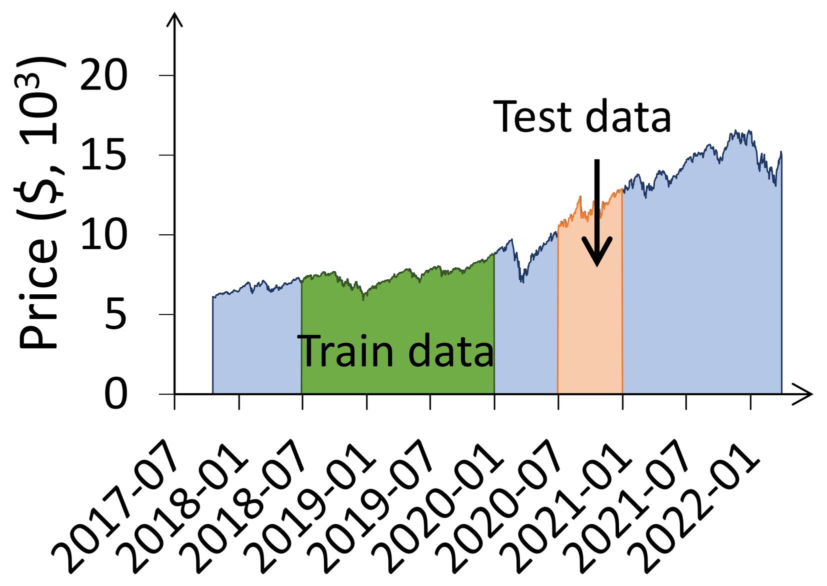
2. Model architecture



3. Experimental setup

Dataset

- US stock market data (NASDAQ, NYSE, AMEX)
- Data from January 2018 to March 2022
- 3,086 financial assets
- Knowledge graph
  - Extracted from Wikidata
  - 102,739 entities
  - 457,758 relations
- Temporal split



Procedure

- Predict ROI at six months in the future.
- Identify top-10 predicted assets.
- Compute average 6 months ROI at the top 10 over the different dates (Mondays in test period).

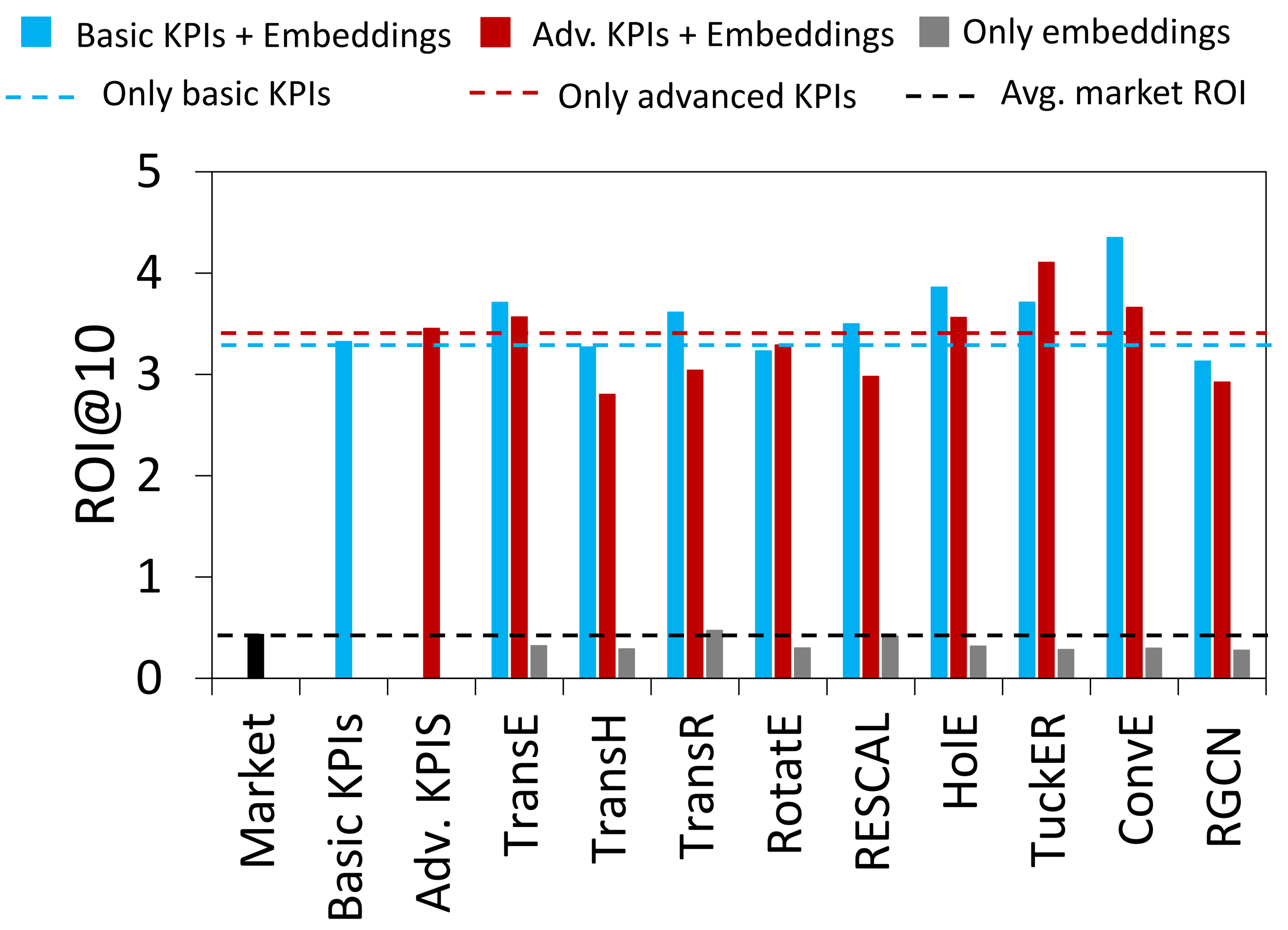
Technical indicators (KPIs)

- Basic
  - Average price
  - Return on investment
  - Volatility
- Advanced:
  - Average true range
  - Momentum
  - Relative strength index
  - Chaikin oscillator
  - Minimum and maximum price
  - …
  - \+ all basic indicators

Knowledge graph embeddings

- Translation-based
  - Relations are defined as translations between two entities in a vector space.
  - **Methods:** TransE, TransH, TransR, RotatE
- Semantic-information based
  - Match latent semantics of entity and relation embeddings.
  - **Methods:** RESCAL, HolE, TuckER
- Neural network-based
  - Use graph convolutional networks to propagate neighborhood information for computing the embeddings.
  - **Methods:** ConvE, RGCN

4. Results & Conclusions



- Basic and advanced KPI models (only technical indicators) beat the market.
- Methods using only embeddings do not improve the market, as they do not use any temporal information.
- Embedding models can improve the effectiveness of our model:
  - TransE, TuckER, HolE and ConvE beat the baselines with both basic and advanced embeddings.
  - Both variants of TuckER with KPIs obtain a statistically significant advantage over the baselines (pairwise Wilcoxon test, $p<0.05$).
  - ConvE + basic KPIS also improves both baselines significantly, obtaining the best ROI@10 at six months (4.3496).

- **Conclusion**
  Knowledge graph embeddings can help identifying top profitable assets when used along technical indicators

HORIZON 2020

This work was carried out as part of the Infinitech project which is supported by the European Union's Horizon 2020 Research and Innovation programme under grant agreement no. 856632.