# Risk classification methods in Robot Advisory platforms

Paolo Giudici[1] Gloria Polinesi[2]

[1]Universitá di Pavia [2]Universitá di Ancona

June 6, 2018

# Robo-Advisors: definition

"Applications that combine digital interfaces and algorithms, and can also include machine learning, in order to provide services, ranging from automated financial recommendations, to contract brokering, to portfolio management to their clients. Such advisors may be standalone firms and platforms, or can be in-house applications of financial instiutions." (FSB, 2017)

# Robo-Advisors: operations

Robo-advisors build personalised portfolios for investors, on the basis of their risk tolerance and aversion, net income, age and family status, possibly collected through on-line questionnaires.

Robo-advisors typically operate algorithmic portfolio allocation adjust portfolio weights "ex-post", to take into account investors' risk preferences

# Research questions

1. Do robo-advisory portfolio allocation match investors' risk preferences?

2. Can we build an integrated portfolio allocation algorithm, that incorporates investors' risk preferences into the model?

# Proposal

1. Descriptive approach: build clusters of assets using network models, and check if they match investors' risk profile clusters
2. Predictive approach: build allocation algorithms using neural networks, using investors' risk profile clusters

# Descriptive analysis: background

Mantegna (1999) suggests to employ the correlation matrix to detect the clustering structure present among the assets of a financial market.

The application of hierarchical clustering techniques, such as the Minimum Spanning Tree (MST) to the correlation matrix reveals the existence of common behaviour of groups of asset returns: such groups identify communities that can correspond to economic sectors or subsectors.

# Minimum Spanning Trees

The Minimum Spanning Tree (MST) is a graph tree that allows to shrink links connecting asset returns from the $\frac{N(N-1)}{2}$ present in the empirical correlation matrix to $N-1$.

It is the minimal shortest path in terms of sum of distances between all assets.

`https://en.wikipedia.org/wiki/Kruskal%27s_algorithm#/media/File:MST_kruskal_en.gif`

# Random Matrix Theory

Plerou (2002) demonstrate that the eigenvalues deviating from a random matrix well convey the meaningful information stored in the correlation matrix.

Accordingly, a 'filtered' correlation matrix can be constructed retaining only the deviating eigenvectors

RMT can improve MST community detection.

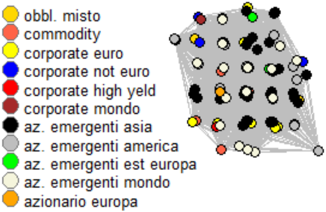
# Data

We were provided data by two European Robo-Advisors, which includes:

- Daily closing price of 450 Exchange Traded Funds (ETFs) in the time period from January 2006 to February 2018;
- Responses to a risk profile questionnaire by 139 anonymous investors, consisiting of 14 questions, each with 4 or 5 possible responses (from the least to the most risky)

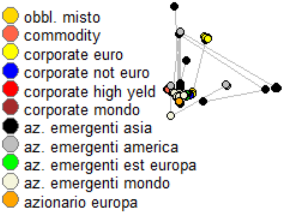
# Complete graph ($\frac{N(N-1)}{2}$ links)

$$N = 92 \quad T = 200$$




Classification of ETFs from Borsa Italiana.

# MST ($N-1$ links)




$N = 92 \;\; T = 200$

Classification of ETFs from Borsa Italiana.

# RMT + MST ($N - 1$ links)

$$N = 92 \quad T = 200$$


Classification of ETFs from Borsa Italiana.

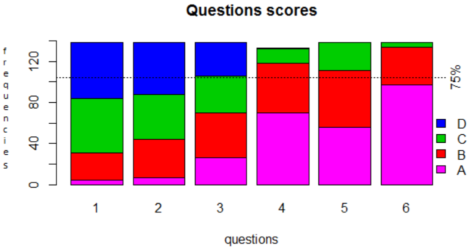
# Predictive analysis: background

1. Scherer (2016) classifies investors based on their risk profile
2. We also classify investors using the MIFID questionnaires
3. The obtained clusters will be used as response variable in a feed-forward neural network.

# Discriminant variables: Scherer's

1. objective
2. time horizon
3. amount of the investment
4. risk tolerance
5. type of investments already done
6. yearly income

# Discriminant variables: ours




Questions 4,5,6 show for over 75% responses A and B. Indeed they are not discriminant.

# The obtained clusters

Based on the first three questions, the optimal number of risk profile classes is equal to 4.

| Risk class | Proposal | Robot Advisor | Match frequencies |
|---|---|---|---|
| I | 13 | 21 | 9 |
| II | 33 | 38 | 19 |
| III | 43 | 36 | 22 |
| IV | 50 | 44 | 41 |
| Sum | 139 | 139 | 91 (65%) |

Table: Frequencies in each risk class, comparing our proposal and what done by the Robot Advisor.

# The neural network model - I

- We consider each ETF return as an input variable, and the cluster group as the output variable
- The neural network is fed with the daily multivariate returns, and their cluster group, in a monthly holding period
- We can thus learn the "allocation" rule of each ETF. The more volatile days have the highest importance weights.
- We can then use the network to predict the actual risk class for a new ETF, and for any actual investor's portfolio, and compare it with the expected risk profile.

# The neural network model - II