HANDS-ON MACHINE LEARNING FOR TACTICAL ASSET ALLOCATION

Guest Lecture InvestSuite 2021-11-25

GHENT UNIVERSITY

FACULTY OF ECONOMICS AND BUSINESS ADMINISTRATION

InvestSuite

GHENT UNIVERSITY

CONTENT

- INTRODUCTION
- INVESTSUITE
- TAA: TIMING THE MARKET
- MACHINE LEARNING
- HANDS-ON EXAMPLE

InvestSuite

FACULTY OF ECONOMICS AND BUSINESS ADMINISTRATION

INTRODUCTION

ROBO-ADVISORY: QUE?

- **Automated investment advice**

- Prevalent investing climate (of low-interest rate regime, digital-natives starting to work and invest) turned many savers into investors (Grealish & Kolm, 2021)

- Reference to article "Robo-Advisors Today and Tomorrow: Investment Advice Is Just an App Away."

ROBO-ADVISORY: FROM INVESTING PRINCIPLES AND ALGORITHMS TO FUTURE DEVELOPMENTS

Chapter to appear in the book
"Machine Learning in Financial Markets: A Guide to Contemporary Practice"
Cambridge University Press, 2021. Edited by A. Capponi & C. A. Lehalle

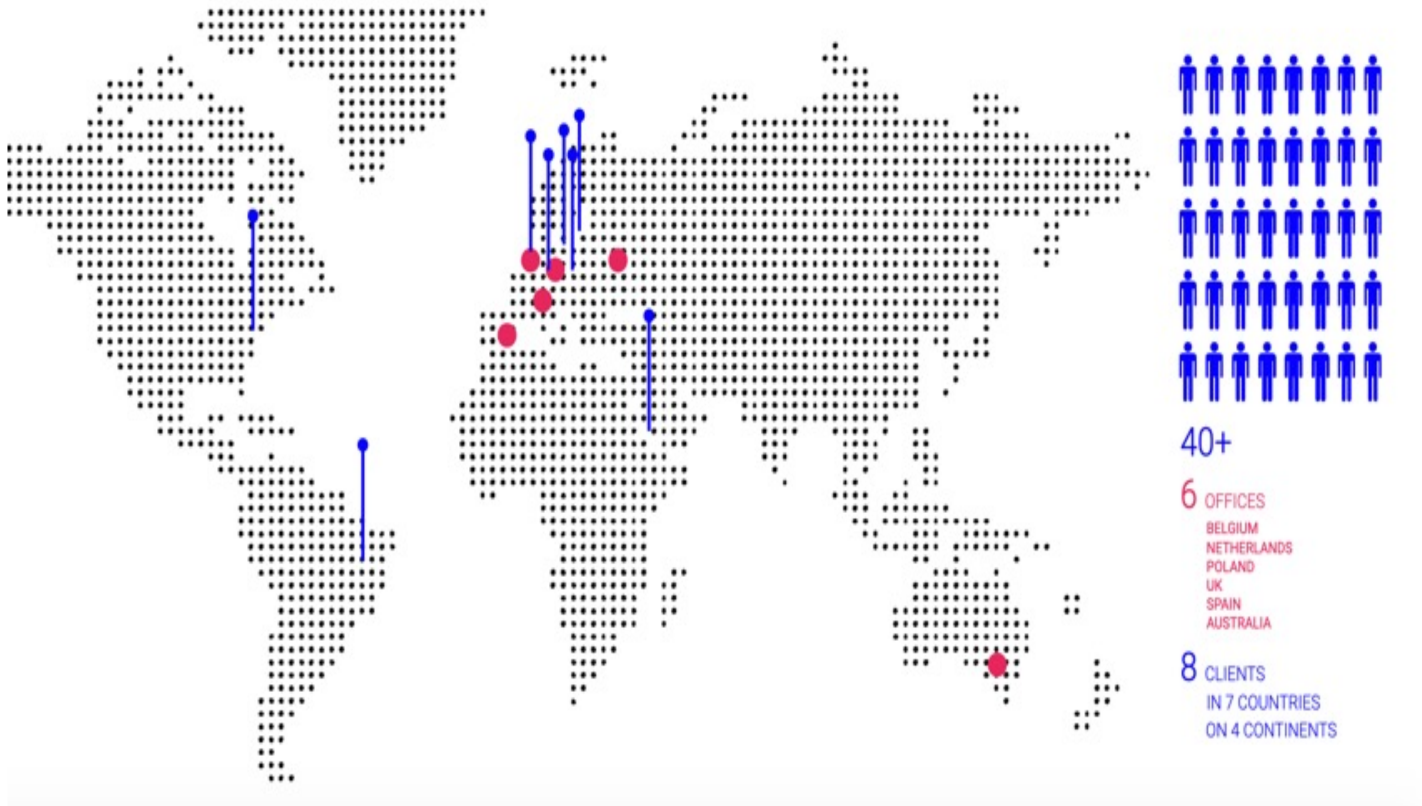
ADAM GREALISH[a] AND PETTER N. KOLM[b]

ABSTRACT. Advances in financial technology have led to the development of easy-to-use online platforms referred to as *robo-advisors* or *digital-advisors*, offering automated investment and portfolio management services to retail investors. By leveraging algorithms embodying well-established investment principles and the availability of exchange traded funds (ETFs) and liquid securities in different asset classes, robo-advisors automatically manage client portfolios that deliver similar or better investment performance at a lower cost as compared to traditional financial retail services.

ROBO-ADVISORY: INVESTSUITE

Key facts

- B2B Robo-Advisory based in **Leuven**
- Founded in **2018**
- Since then grew to a team of **50+** people
- **8 clients** in **7 countries** on **4 continents**
- **4 products** in our current suite
- https://www.investsuite.com/



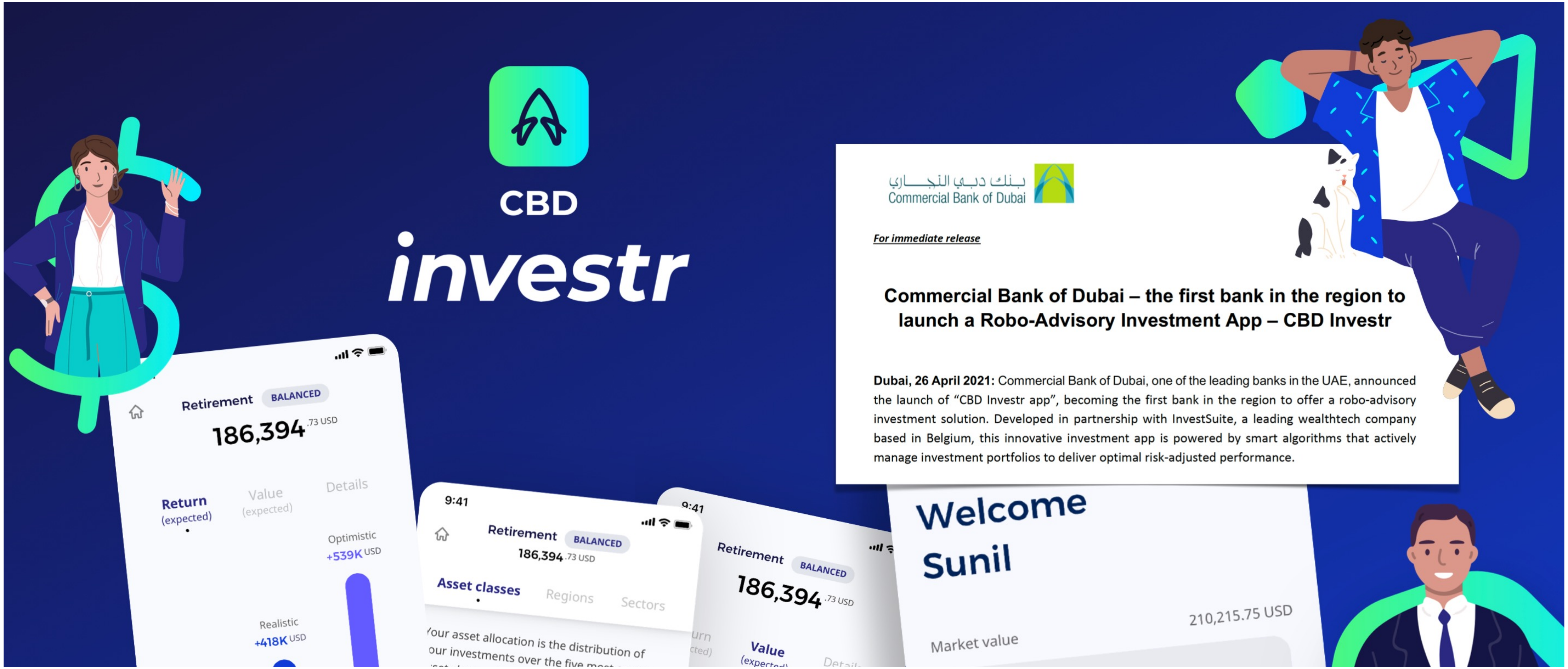
ROBO ADVISOR
A low-cost, customisable digital wealth management tool that converts savings into profitable investment assets

SELF INVESTOR
A white-label execution only platform for easy investing

PORTFOLIO OPTIMIZER
The next-generation quant tools that provide cost-effective solutions for more efficient portfolio management

STORYTELLER
A worldwide 'first' new way of 'telling the story' of retail clients' portfolio performance

ROBO-ADVISORY: INVESTSUITE



ROBO-ADVISORY: PORTFOLIO CONSTRUCTION

- Portfolio construction in practice

- Large gap between academia and existing robo solutions

- Beketov (2018) investigated **219 robo-advisors** covering the vast majority of market players (including Wealthfront, Betterment, etc.)

- Reference to article

Source: Beketov 2018

ROBO-ADVISORY: PORTFOLIO CONSTRUCTION

Table 3 Occurrence of different methodological frameworks within the Robo Advisors analyzed

Methodological framework	Occurrence (%)
Modern Portfolio Theory	39.7
Sample Portfolios	27.4
Constant Portfolio Weights	13.7
Factor Investing	2.7
Liability-Driven Investing	2.7
Risk Parity	1.4
Full-Scale Optimization	1.4
Constant Proportion Portfolio Insurance	1.4
Mean Reversion Trading	1.4
Other	8.2

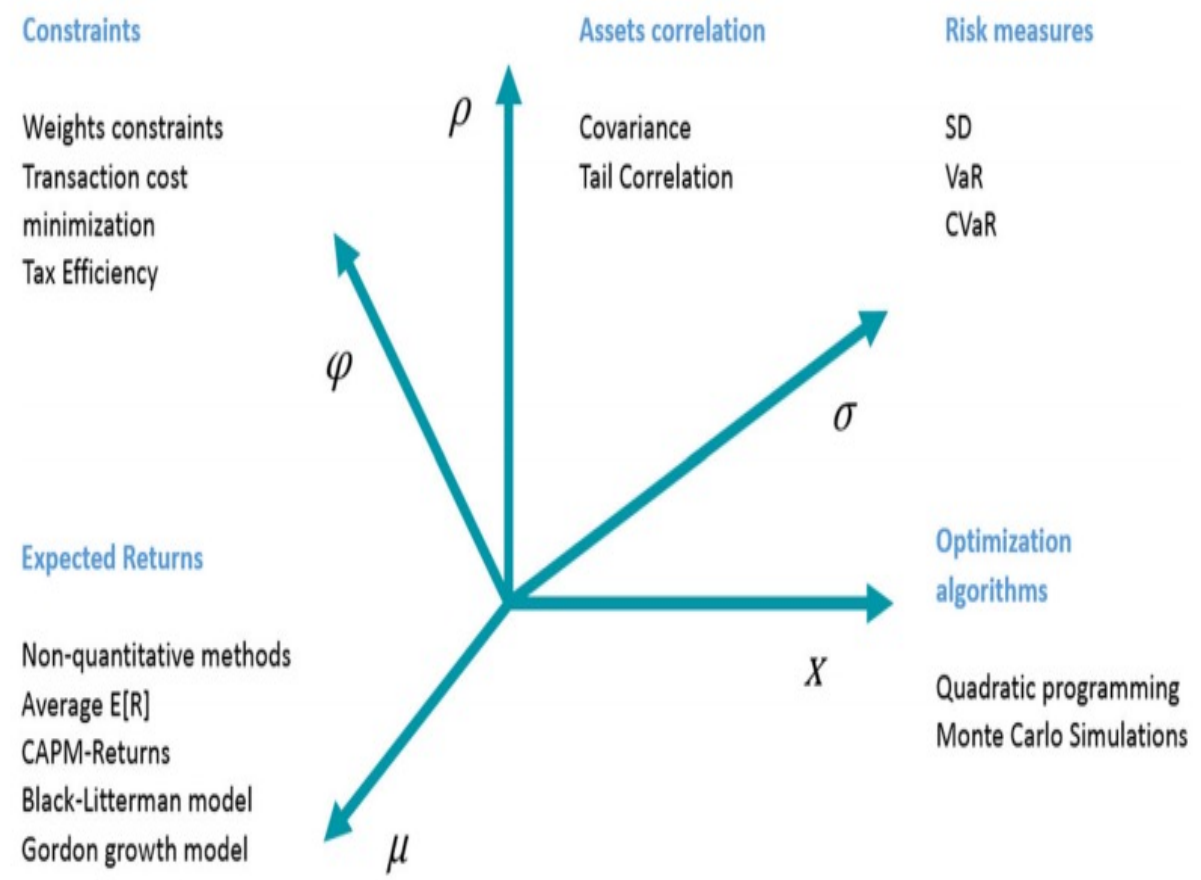
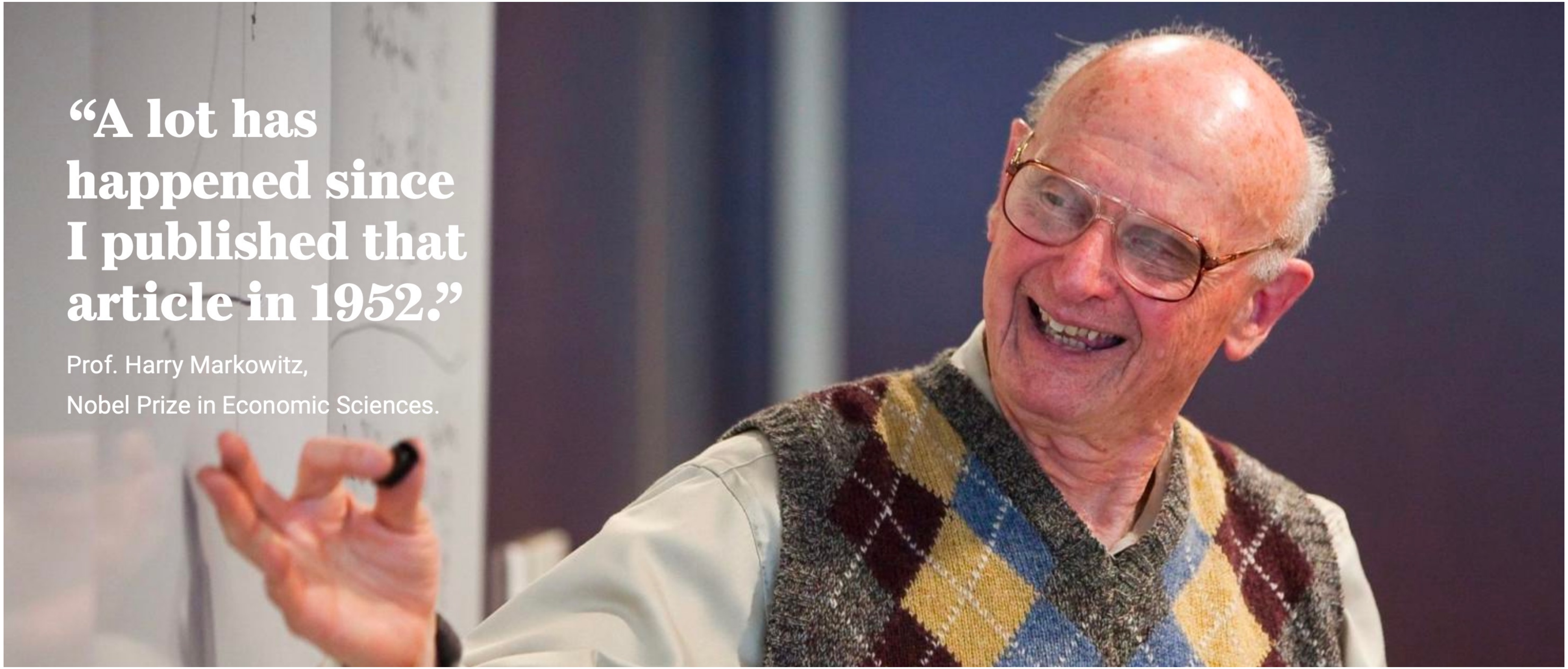


Fig. 3 Schematic of the "Multidimensional improvement of Modern Portfolio Theory" in Robo Advisors. The methods mentioned are those that used in RAs, and they do not comprise a comprehensive list of the methods that are or can be used to improve the Modern Portfolio Theory framework in general. Note: VaR and CVaR optimization are frequently considered to be alternatives to Modern Portfolio Theory rather than improvements to this framework

Source: Beketov 2018

ROBO-ADVISORY: PORTFOLIO CONSTRUCTION



ROBO-ADVISORY: IVAR

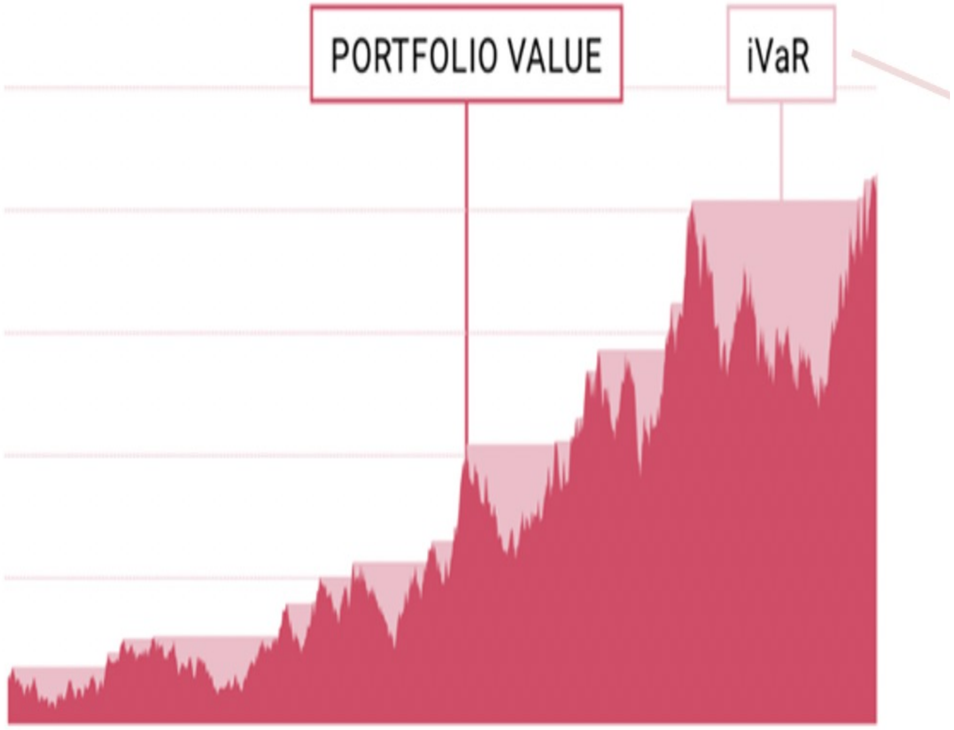
- **Premise**: Attractive alternatives to the savings account

- Returns are hard (impossible ?) to predict, and ML is not an oracle

- Risk is easier to predict, hence traditional volatility models and risk-based portfolio optimization

- Essentially a more **persistent** feature of financial time series, for which it's more natural to optimize

- **Integrated value-at-risk ("iVaR")** is InvestSuite's proprietary risk framework that aims at reducing (1) the **size** and (2) the **frequency** of **losses**, as well as (3) the **time** to recoup them.



$$\begin{aligned} \min_w \quad & \mathbb{E}(\xi(w)) \\ \text{s.t.} \quad & \xi = m_t - w\boldsymbol{\Pi}_t \\ & m_t \geq m_{t-1} \\ & w\mathbf{I}^N = 1 \end{aligned}$$

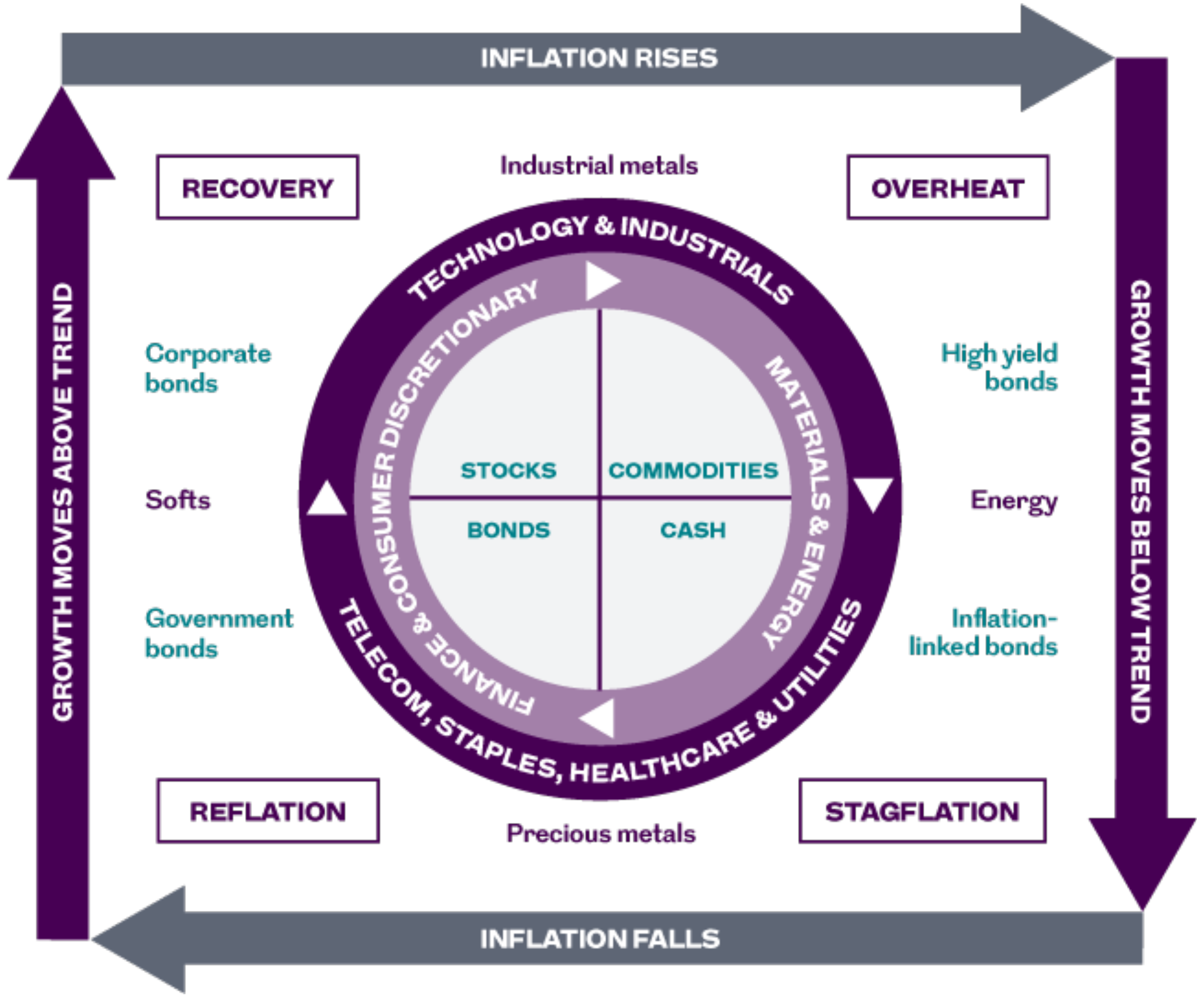
TACTICAL ASSET ALLOCATION:
TIMING THE MARKET

GHENT UNIVERSITY

FACULTY OF ECONOMICS AND BUSINESS ADMINISTRATION

TACTICAL ASSET ALLOCATION

- **Timing the market**: what time is it on the investment clock?
- In essence, **beta < 1** vs. **beta > 1** strategies based on macro-economic indicators or trend detectors.
- 'Risk-on, risk-off' (cf. application) or less binary strategies.

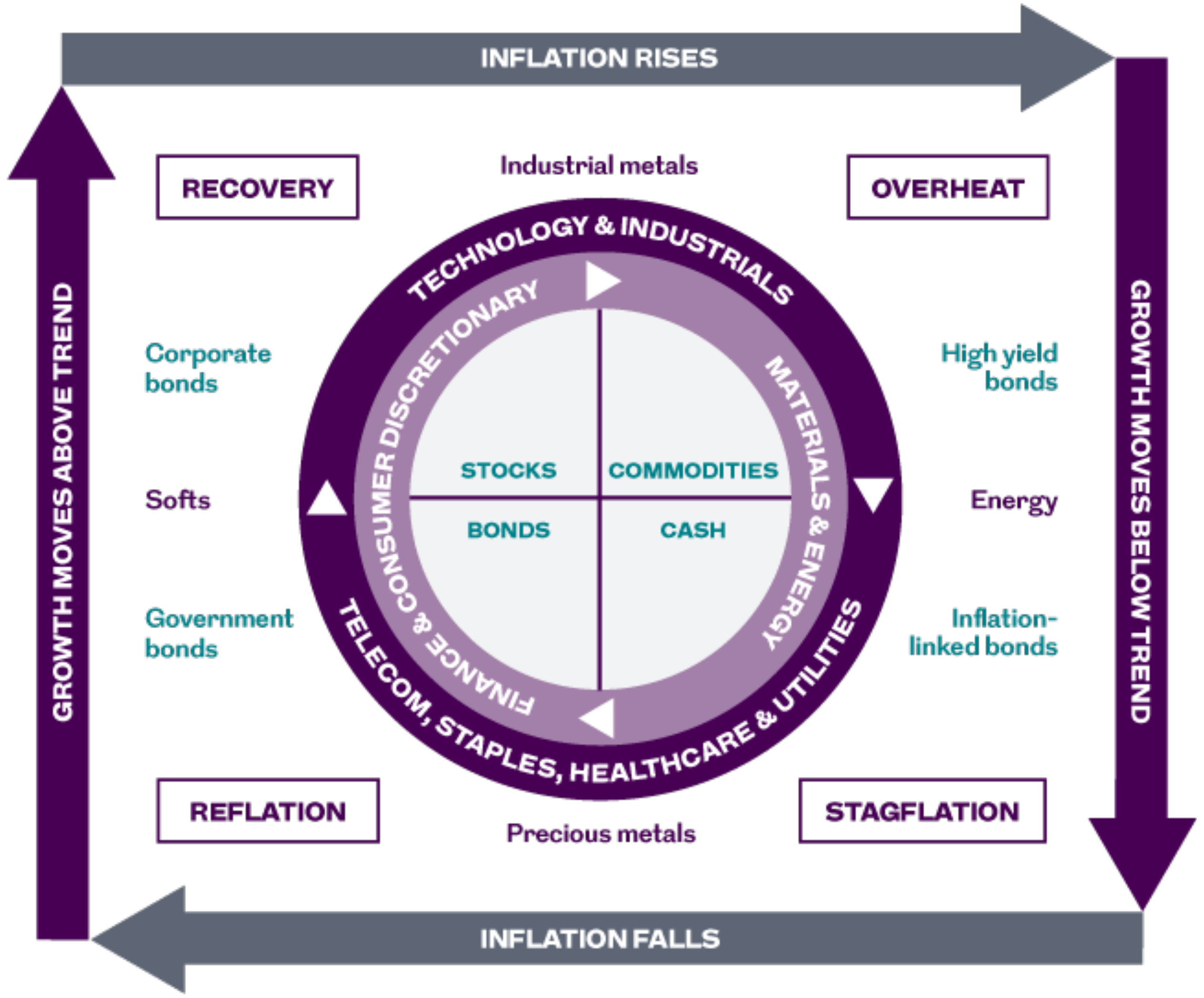


Merill Lynch Investment Clock
(from: Greetham, Trevor, and H. Hartnett. "The Investment Clock Special Report# 1: Making Money From Macro," Merrill Lynch, 2004.)

GHENT UNIVERSITY
FACULTY OF ECONOMICS AND BUSINESS ADMINISTRATION

TACTICAL ASSET ALLOCATION

- (Early) warnings for overheating…
  - Valuation based: Shiller index, PE ratios (e.g. Shiller 2021)
  - Cointegration-based tests and real-time Bubble detectors (e.g. Philips, Shi, Wu)
- …and relative performance:
  - Relative Strength indices (e.g. SMA indicators)
  - Trend-based methods (e.g. aggregated/index momentum signals, e.g. Schnetzer 2020)



Merill Lynch Investment Clock

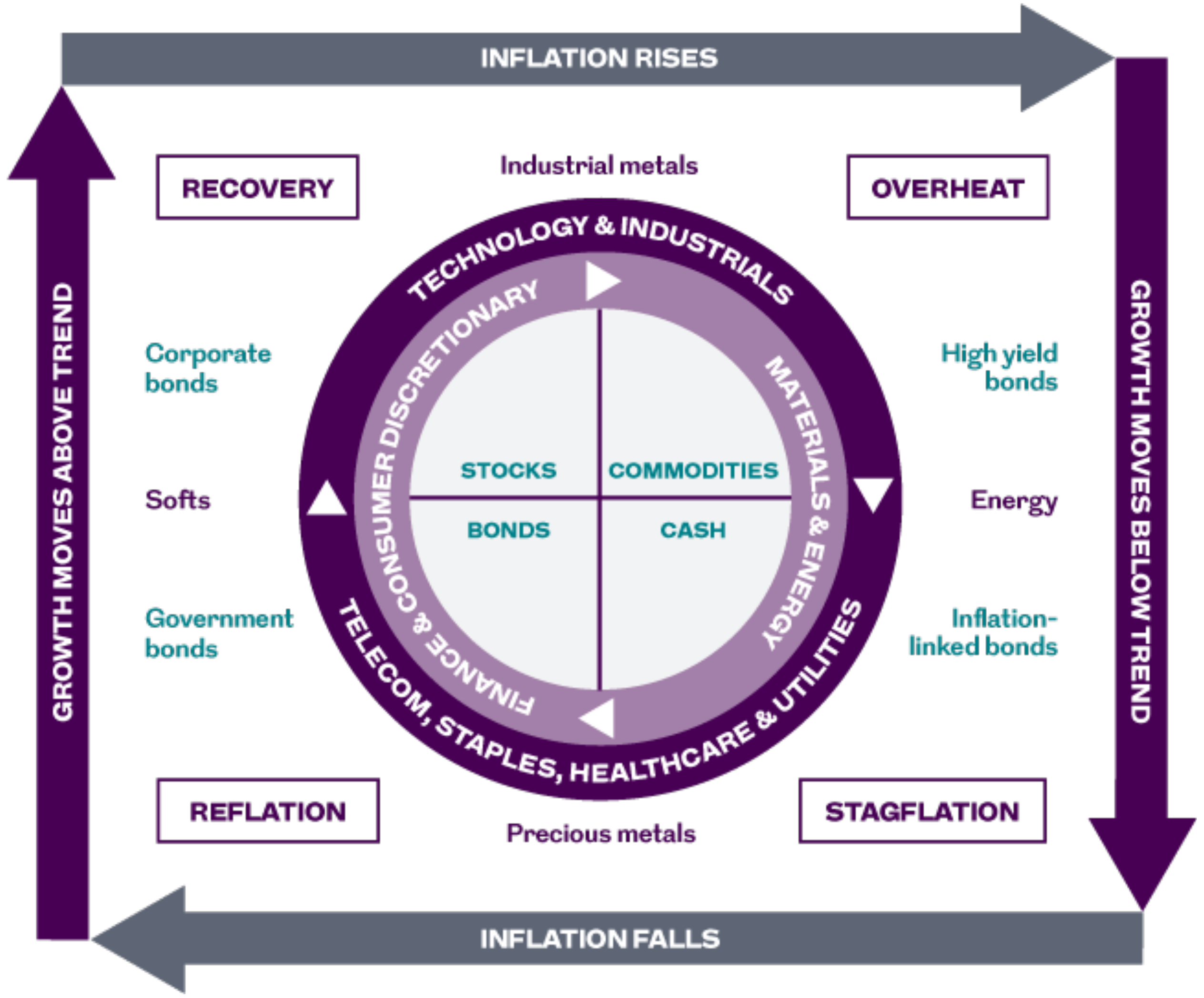
(from: Greetham, Trevor, and H. Hartnett. "The Investment Clock Special Report# 1: Making Money From Macro," Merrill Lynch, 2004.)

GHENT UNIVERSITY

FACULTY OF ECONOMICS AND BUSINESS ADMINISTRATION

TACTICAL ASSET ALLOCATION

- **In sum: Gauging the market regime…**
  - Regime-based asset allocation far from trivial, but let's look at a **simple** ML example based on the Achilles heel of traditional portfolio construction: **the correlation matrix**.
  - Let us focus on **ML** and **feature engineering** (in not too domain-specific terms).



Merill Lynch Investment Clock

(from: Greetham, Trevor, and H. Hartnett. "The Investment Clock Special Report# 1: Making Money From Macro," Merrill Lynch, 2004.)

FEATURE ENGINEERING: UNDERSTANDING CLASSICAL MODEL FAILURE

- The classical issue in traditional models (cf. slide on Beketov 2018):

  Correlations break down when needed most…
  …when volatility spikes.

- Well-known issue of mean-variance portfolio construction and correlation-based methods in general.

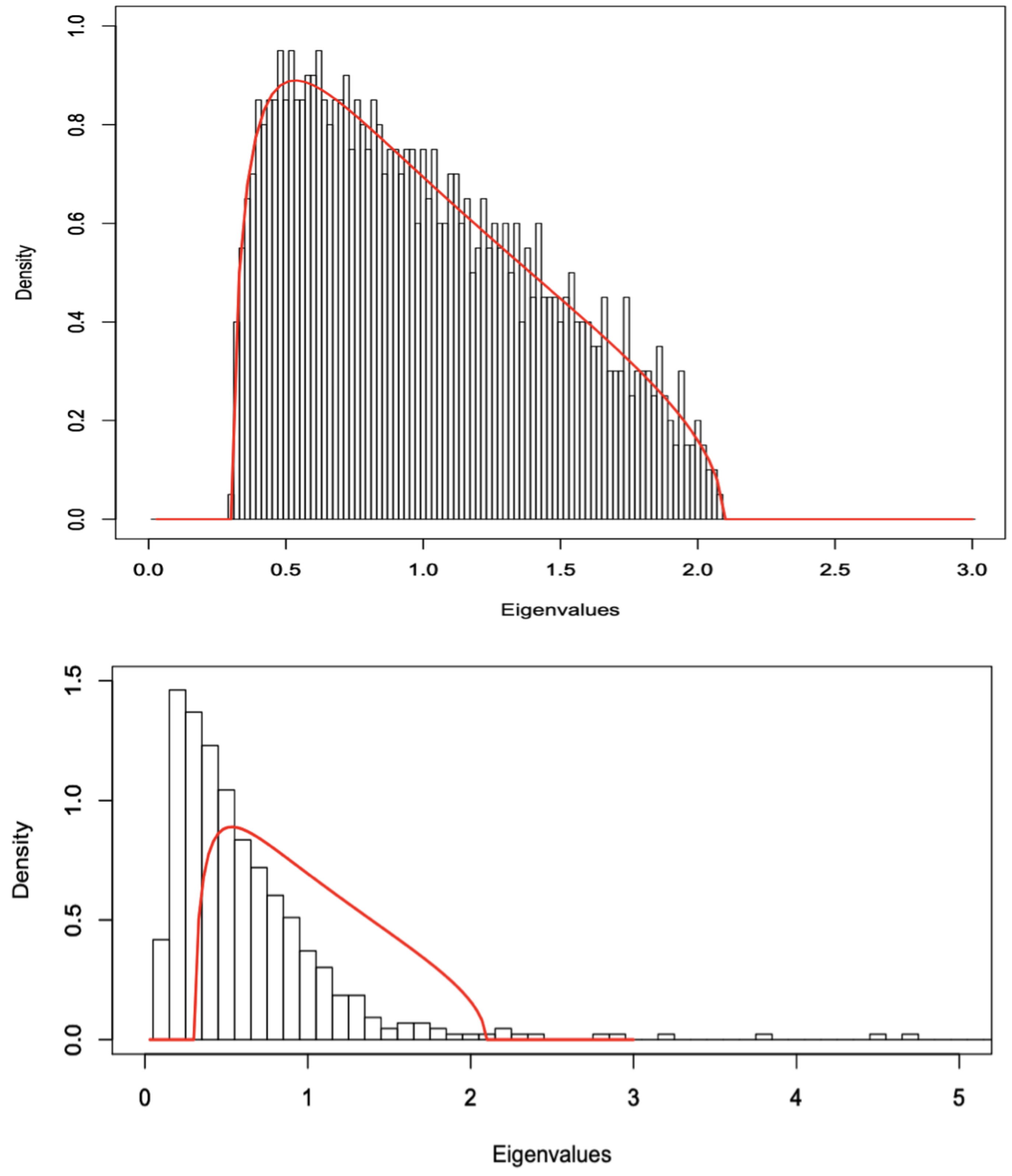
- **How can we turn this around and use the stylized facts to our advantage for regime prediction?**
- (One) recent proposal in literature: thoroughly understand when and why model-driven approaches fail and recognize these incidents early with ML (de Prado 2018, Marti 2021)

STYLIZED FACTS OF FINANCIAL CORRELATION MATRICES

(1) The distribution of financial correlations is significantly shifted to the **positive**, i.e. most assets are positively correlated.

(2) Eigenvalues follow the **Marcenko-Pastur (MP)** distribution, with the exception of the first very large eigenvalue (the market) and a couple of other large eigenvalues (the industries), e.g. PCA.

=> Determines **conditioning** of covariance matrix (condition number = ratio first and last eigenvalue, and measure for precision. The lower the better!)



Density of the eigenvalues of a random vs. financial covariance matrix

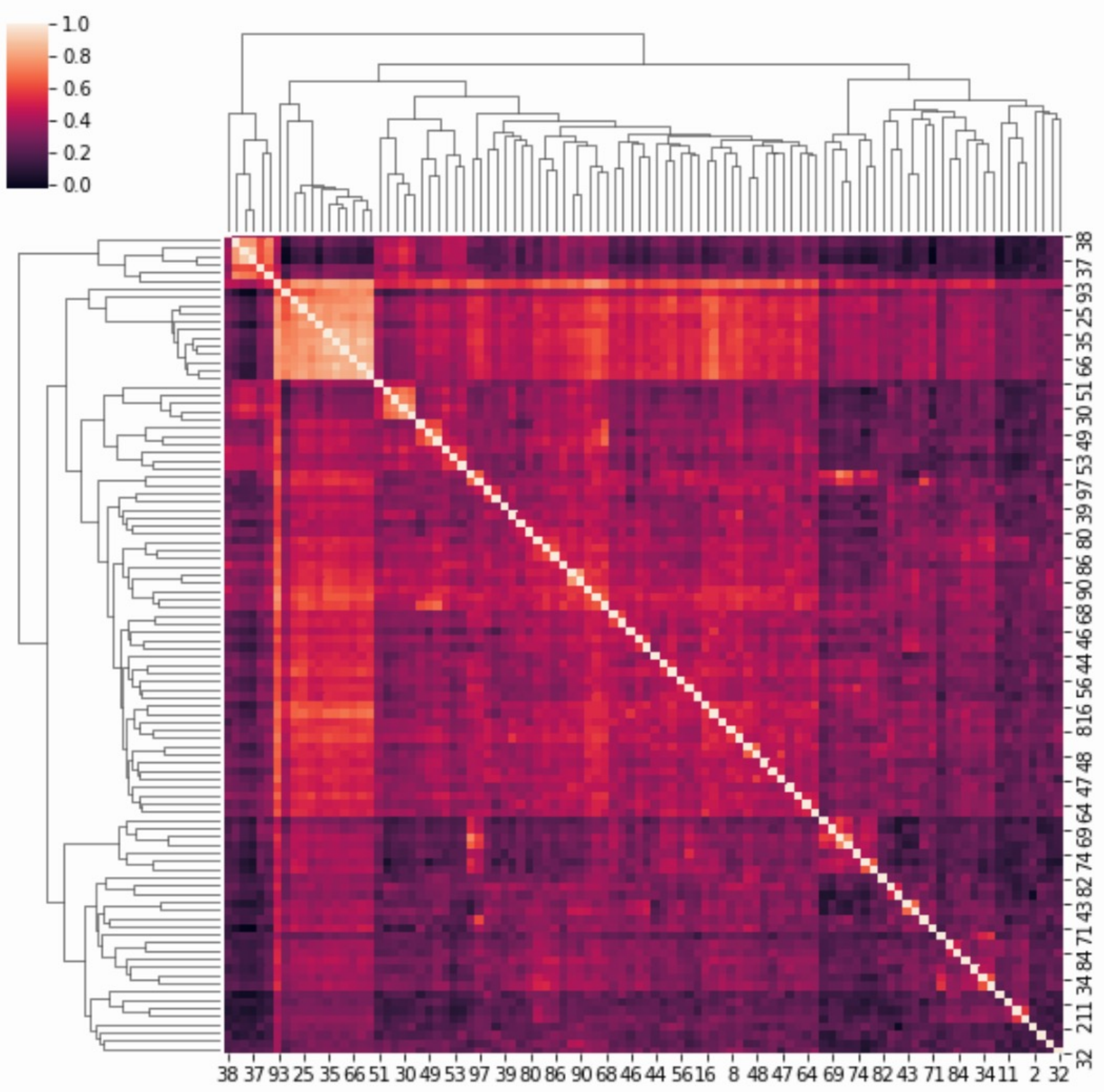
GHENT UNIVERSITY

FACULTY OF ECONOMICS AND BUSINESS ADMINISTRATION

STYLIZED FACTS OF FINANCIAL CORRELATION MATRICES

(3) Perron-Frobius property: the first eigenvector has positive entries, i.e. all assets typically have positive exposure to the market.

(4) Hierarchical structure of correlations.



Correlation heatmap with dendrogram, the brighter the higher the correlation. Higher correlated assets are put next to each other!

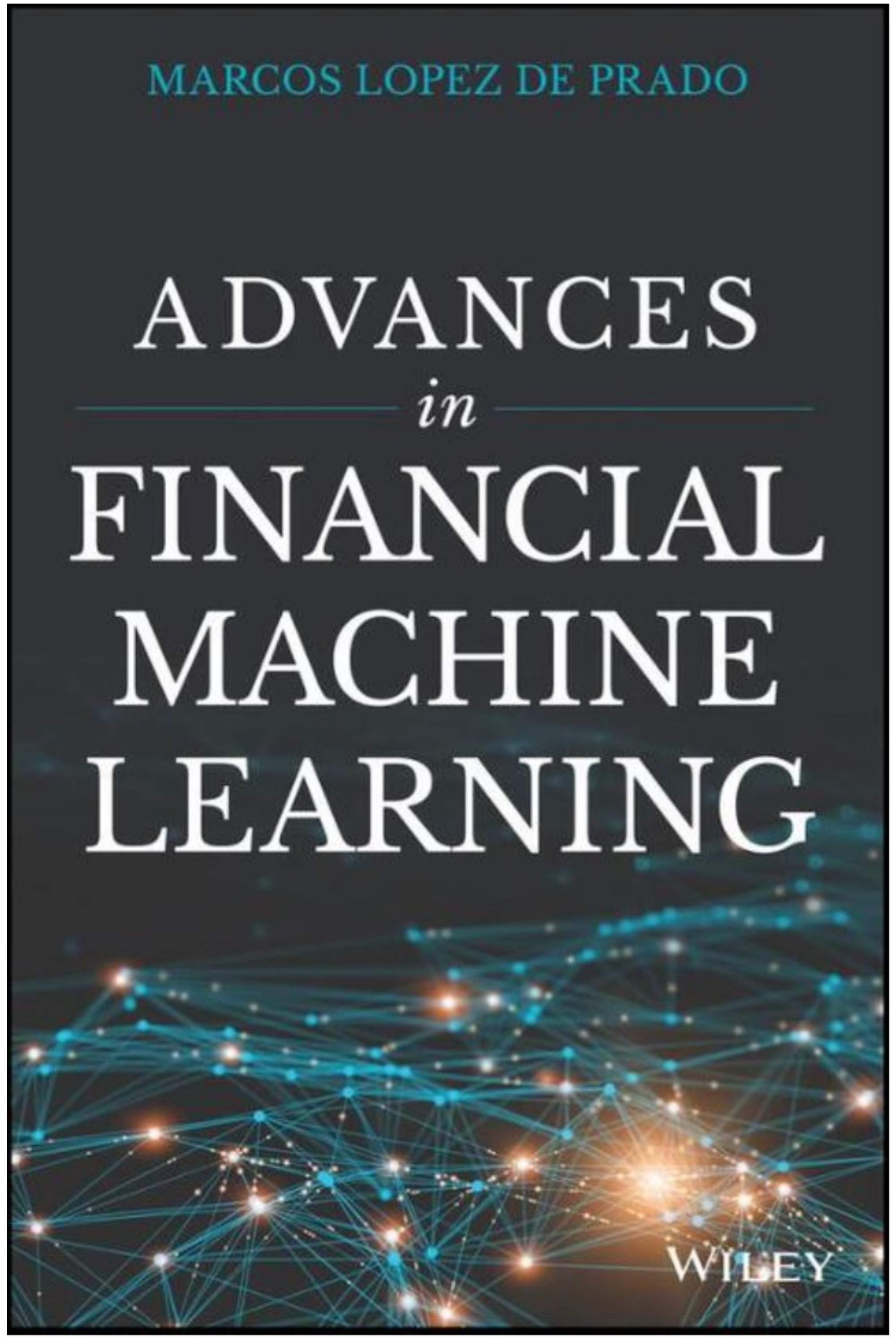
HIERARCHICAL RISK PARITY

- **Curse of Diversification**: precision covariance inversion is inversely proportional to concentration eigenvalues (essentially an ill-posed problem with high degree of multi-collinearity).

- Or in plain terms, the higher the average correlation, the worse the portfolio is actually diversified.

- **HRP**: Risk parity over clusters of assets (iteratively).

- **Intuition**: separate allocation over substitutes versus true alternatives increases confidence.
- **Statistics**: cross-cluster correlation has much lower average than original correlation matrix = can be inverted with much higher precision



APPLICATION

GAUGING THE MARKET REGIME USING CORRELATION MATRIX FEATURES: CONTEXT

- **Aim:** market timing using the features of the empirical correlation matrix solely.

- **Two asset classes (instruments):** 100% equity (a SPDR Eurostoxx50 ETF) and 100% Cash.

- **Every month** we make a **risk-on or risk-off decision** using an ML model (RF).

- Link to application:
  https://drive.google.com/drive/folders/14aPCUsyDSCpAf3ExHT00MxDoBdYyM925?usp=sharing

GHENT UNIVERSITY

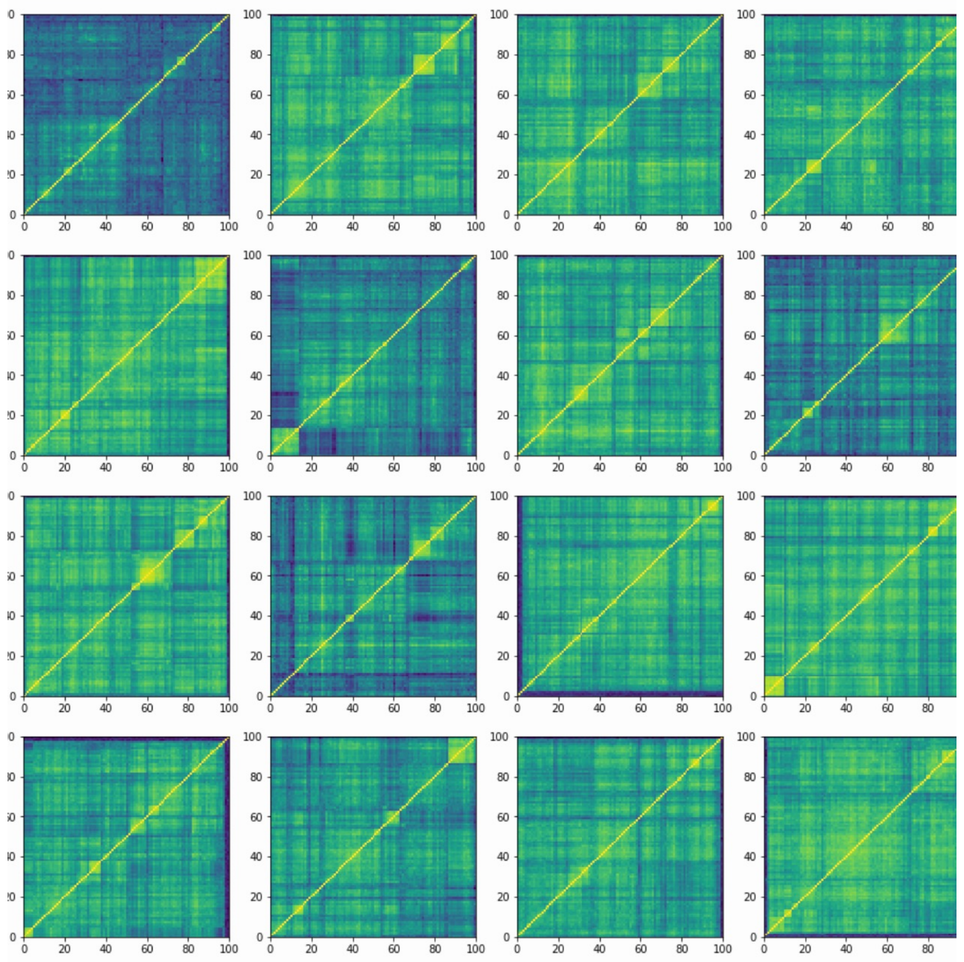
FACULTY OF ECONOMICS AND BUSINESS ADMINISTRATION

GAUGING THE MARKET REGIME USING CORRELATION MATRIX FEATURES: STEPS

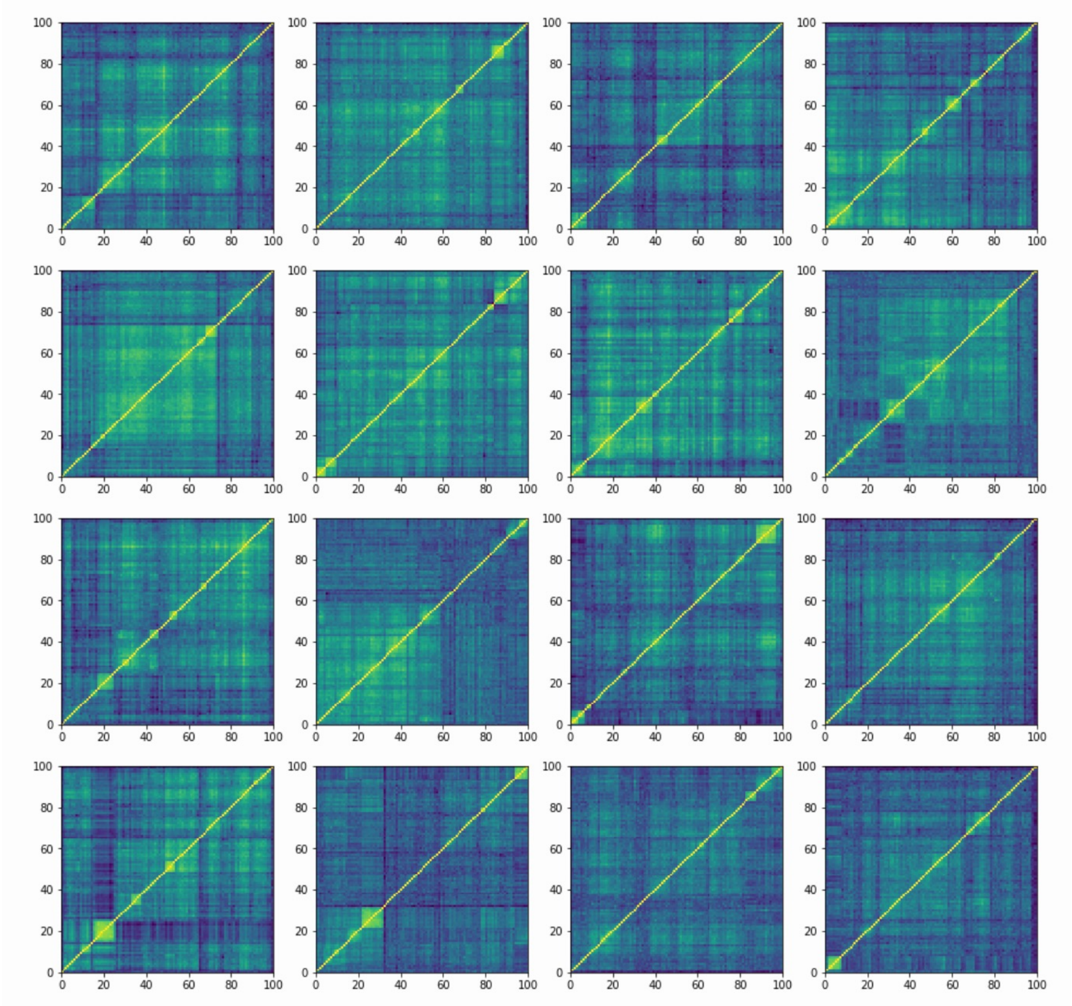
1. **Build intuition** about the features and the labels
2. **Feature engineering** / calculation
3. **Training** the **ML** model
4. **Interpret** the model
5. **Backtesting** the TAA rule
6. **Evaluation** metrics

GAUGING THE MARKET REGIME USING CORRELATION MATRIX FEATURES: (1) INTUITION

RISK MIN - STRESS



RISK MAX - NORMAL



Sample correlation matrices for two different regimes (the brighter the higher the correlation (-1< r < 1))

GAUGING THE MARKET REGIME USING CORRELATION MATRIX FEATURES: (1) INTUITION

- Ref Paper: **cCorrGAN**: Conditional Correlation GAN (Marti 2021).

- Context: researcher wants to use the stylized facts above to predict the market regime and emulate it in a conditional GAN network.
- Created a dataset of 20.000 historical correlation matrices X based on S&P500 data labelled Y = {stress, normal, rally} according to ex-post Sharpe ratio.

- Let us create a similar dataset for **the constituents of the Eurostoxx50** over time.

GAUGING THE MARKET REGIME USING CORRELATION MATRIX FEATURES: (1) INTUITION

- **Intuition**: average correlation is substantially higher during stress, but also the highest correlations, its quantiles, deviations from MP, condition number of the matrix, … etc   => All potential features.

- **Central question: What is a good decision RULE?**

- The author uses **Explainable ML** for this: **RF** classifier and **SHAP** values.

GHENT UNIVERSITY

FACULTY OF ECONOMICS AND BUSINESS ADMINISTRATION

GAUGING THE MARKET REGIME USING CORRELATION MATRIX FEATURES: (2) FEATURE ENGINEERING

- Cf. link to application

- Estimate monthly correlation matrices up until 2017.

- Next we will look at:

  - **Correlation coefficients**
    - Descriptive stats (min, max, mean)
    - Quantiles (1%, 99%, median, 90%, 99-1%, …)

  - **Eigenvalues**
    - Descriptive stats
    - Variance explained
    - Condition number
    - Deviations from MP

```python
# Descriptive statistics
coeffs_stats = coeffs.describe()
for stat in coeffs_stats.index[1:]:
      features[f'coeffs_{stat}'] = coeffs_stats[stat]
# Quantiles
features['coeffs_1%'] = coeffs.quantile(q=0.01)
features['coeffs_99%'] = coeffs.quantile(q=0.99)
features['coeffs_10%'] = coeffs.quantile(q=0.1)
features['coeffs_90%'] = coeffs.quantile(q=0.9)
features['coeffs_99-90'] = features['coeffs_99%'] - features['coeffs_90%']
features['coeffs_10-1'] = features['coeffs_10%'] - features['coeffs_1%']
```

```python
# Concentration of the eigenvalues: variance explained by 1st eigenvalue, top5,
top30, 5ex1, 30ex5, etc.
eigenvals, eigenvecs = np.linalg.eig(model_corr)
features['varex_eig1'] = float(eigenvals[0] / sum(eigenvals))
features['varex_eig_top5'] = (float(sum(eigenvals[:5])) / float(sum(eigenvals)))
features['varex_eig_top30'] = (float(sum(eigenvals[:30])) / float(sum(eigenvals)))
features['varex_5-1'] = (features['varex_eig_top5'] - features['varex_eig1'])
features['varex_30-5'] = (features['varex_eig_top30'] - features['varex_eig_top5'])
```

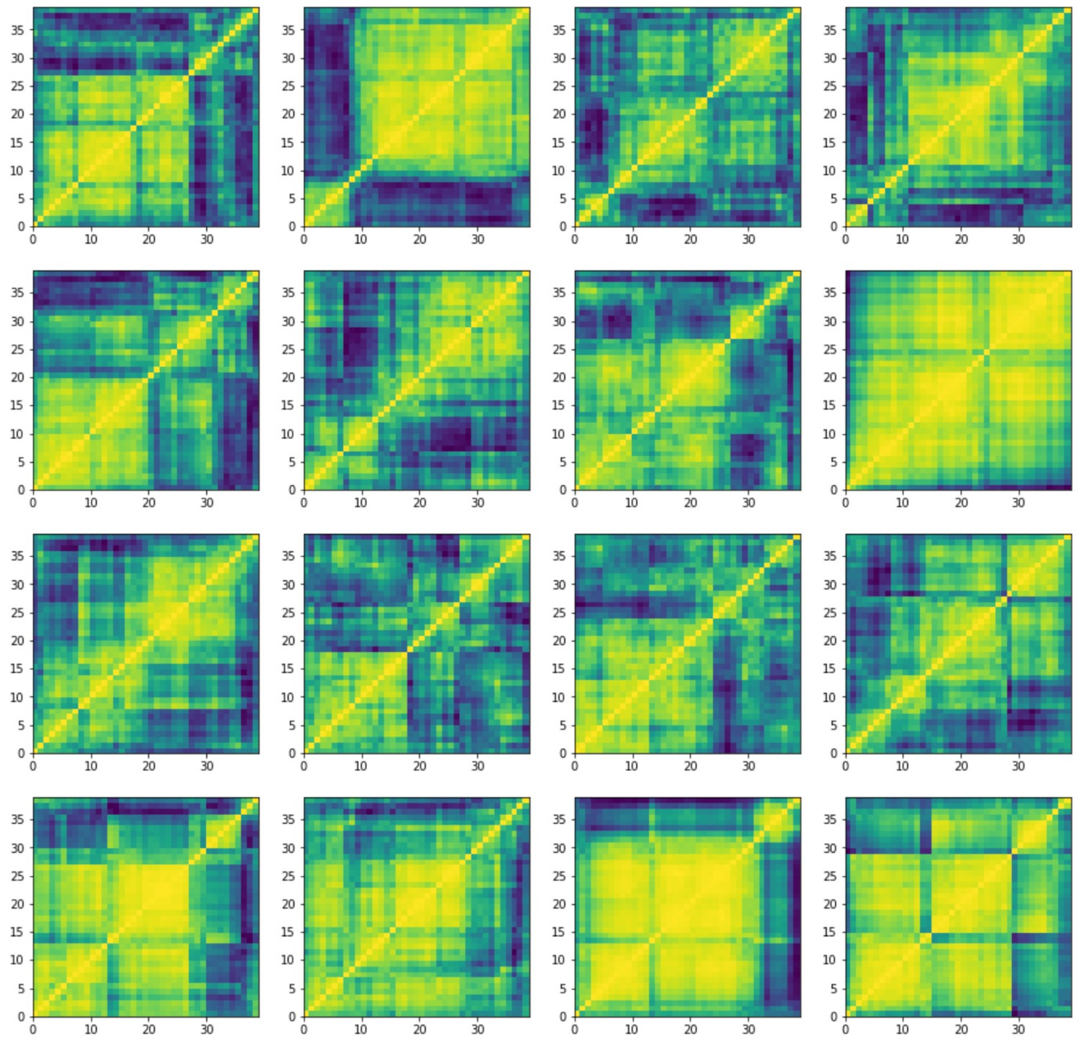
```python
# Variance explained by eigenvals outside of the MP distribution
features['varex_eig_MP'] = (float(sum([e for e in eigenvals if e > MP_cutoff])) / float(sum(eigenvals))
# Condition number
features['condition_number'] = abs(eigenvals[0]) / abs(eigenvals[-1])
```

GHENT UNIVERSITY

FACULTY OF ECONOMICS AND BUSINESS ADMINISTRATION

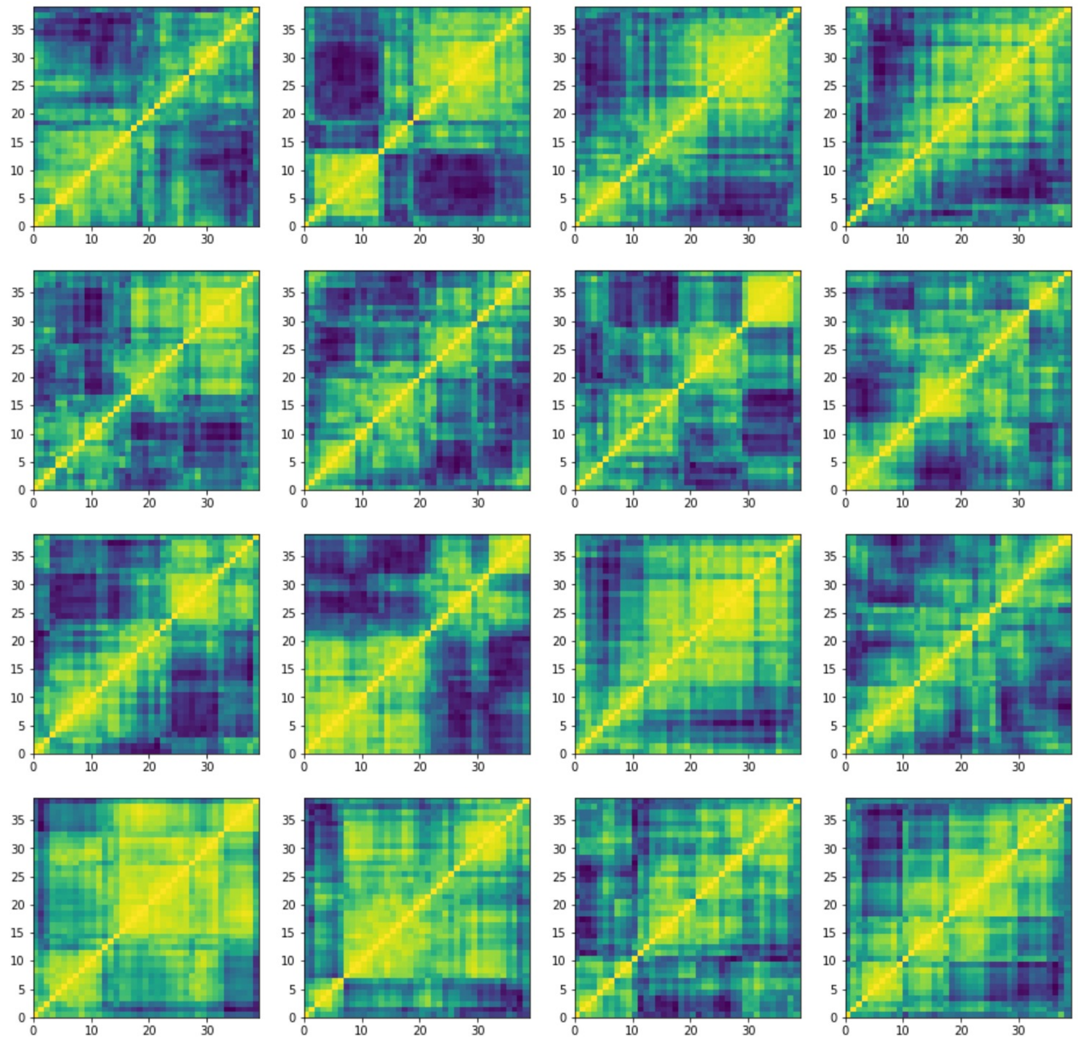
GAUGING THE MARKET REGIME USING CORRELATION MATRIX FEATURES: (2) LABELS

STRESS = 1



Sharpe < -1

NORMAL = 0

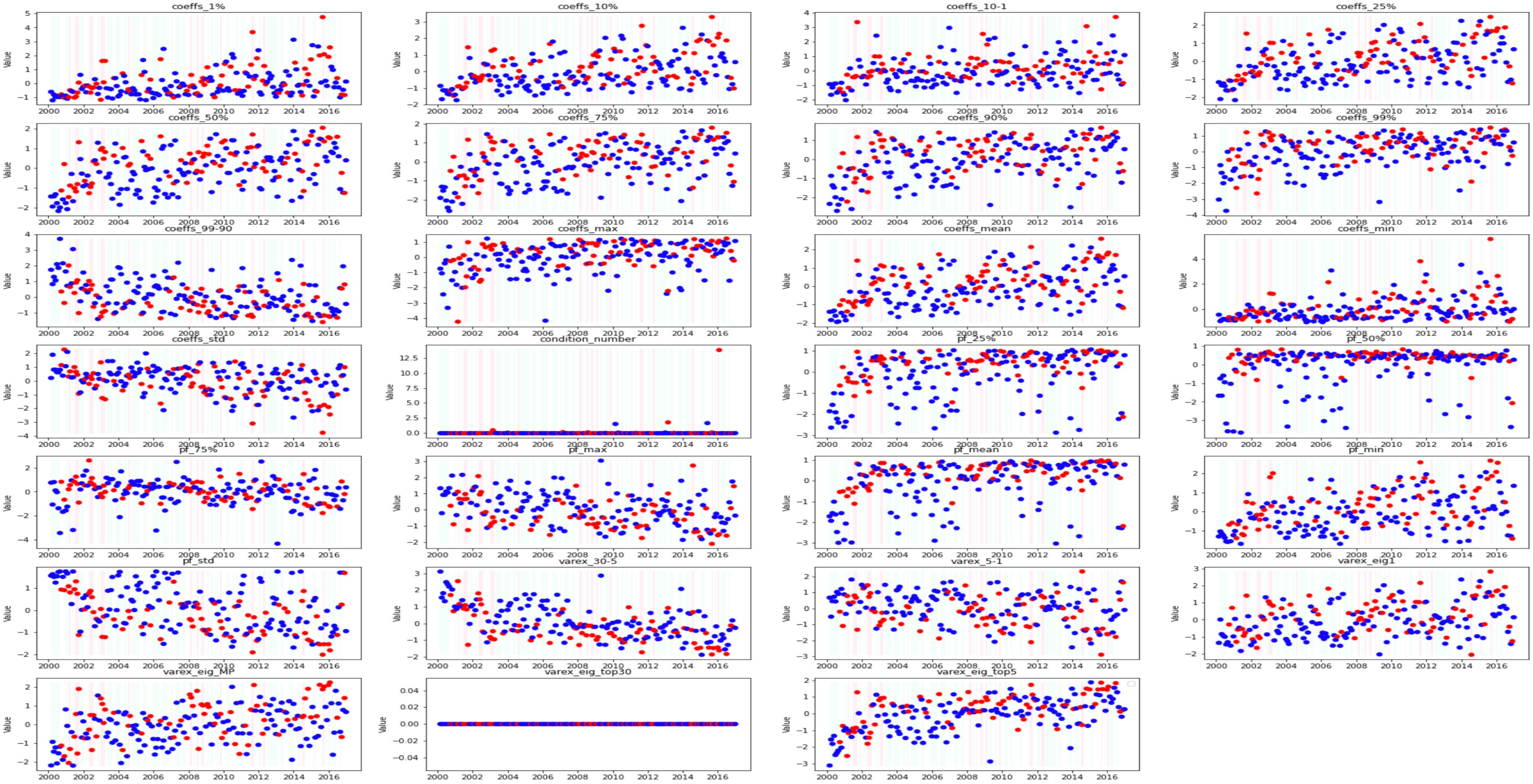


Sharpe > -1

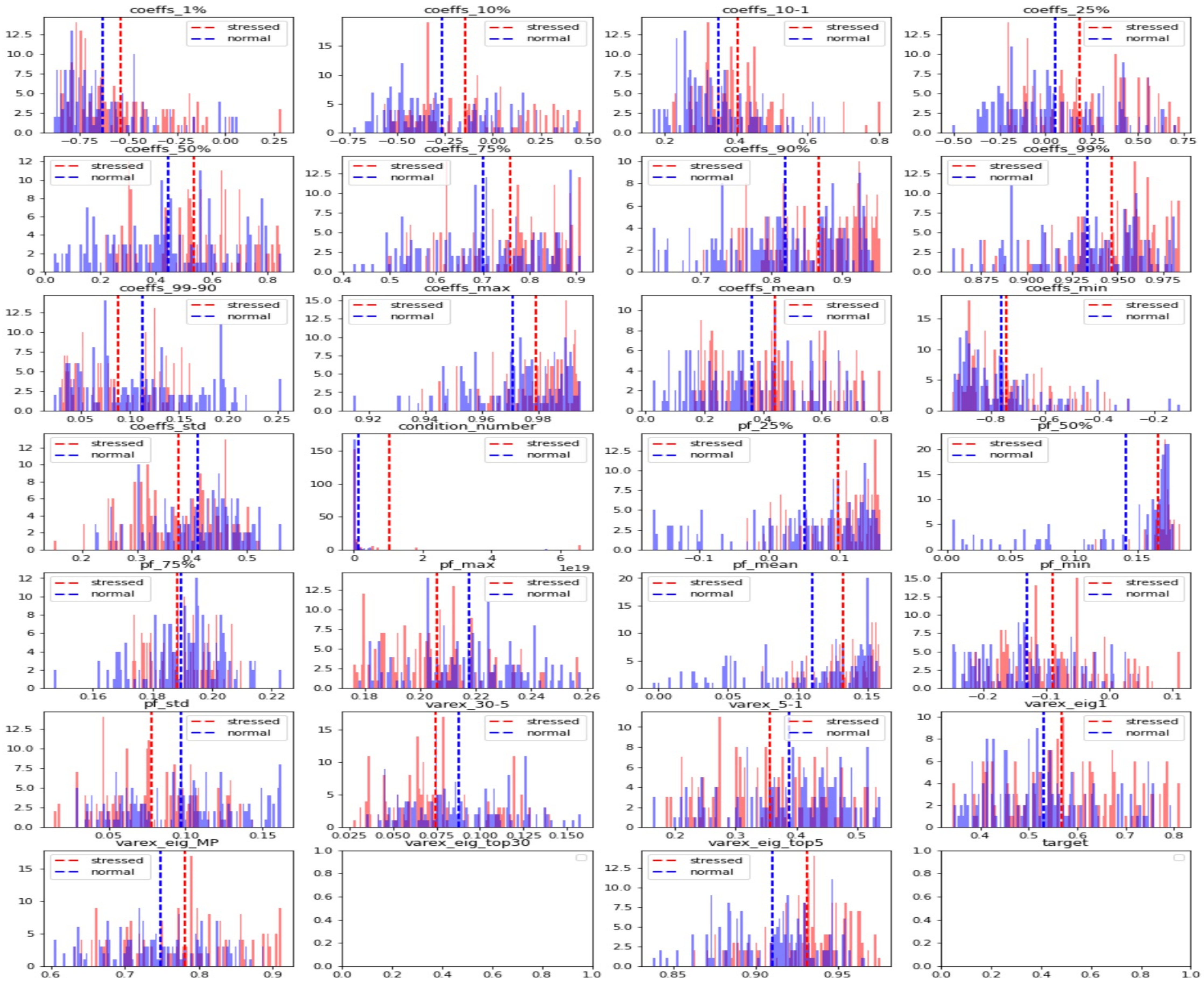
GHENT UNIVERSITY

FACULTY OF ECONOMICS AND BUSINESS ADMINISTRATION

GAUGING THE MARKET REGIME USING CORRELATION MATRIX FEATURES: (2) FEATURE ENGINEERING: EVOLUTION OVER TIME



GAUGING THE MARKET REGIME USING CORRELATION MATRIX FEATURES: (2) FEATURE ENGINEERING, DISTRIBUTION



GHENT UNIVERSITY

GAUGING THE MARKET REGIME USING CORRELATION MATRIX FEATURES: (3) TRAINING THE MODEL

- Simple decision tree (**Random Forest Classifier**) model

```
clf = RandomForestClassifier()
clf.fit(X_train, y_train)

print('Accuracy on train set:', clf.score(X_train, y_train))
print('Accuracy on test set:', clf.score(X_test, y_test))

labels = ['stressed', 'normal']
confusion_mat = confusion_matrix(
y_test, clf.predict(X_test), labels=labels)
```

⇨

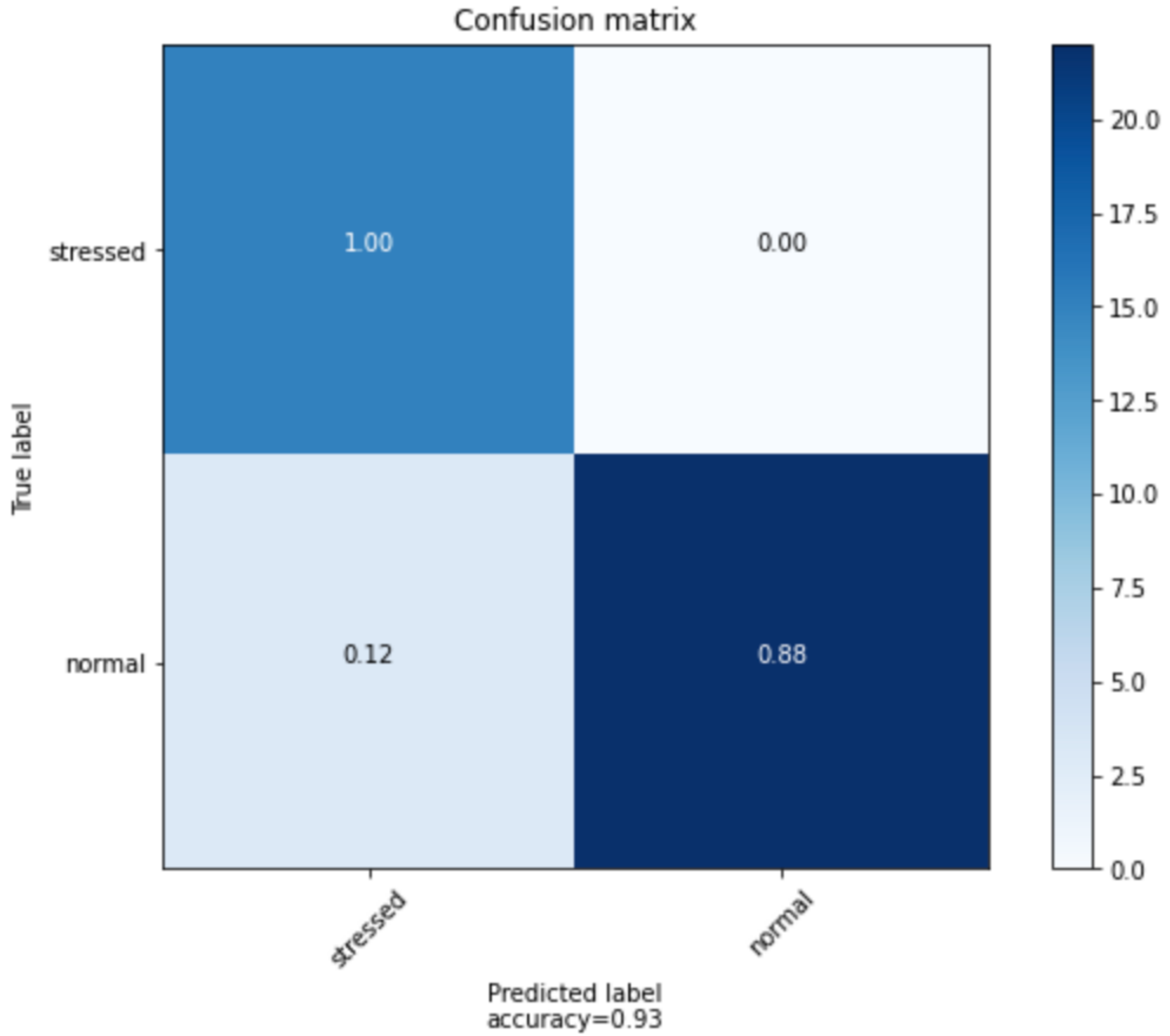
```
Accuracy on train set: 1.0
Accuracy on test set: 0.925
CPU times: user 287 ms,
sys: 4.99 ms, total: 292 ms
Wall time: 293 ms
```

GHENT UNIVERSITY

FACULTY OF ECONOMICS AND BUSINESS ADMINISTRATION

GAUGING THE MARKET REGIME USING CORRELATION MATRIX FEATURES: (4) INTERPRET THE MODEL

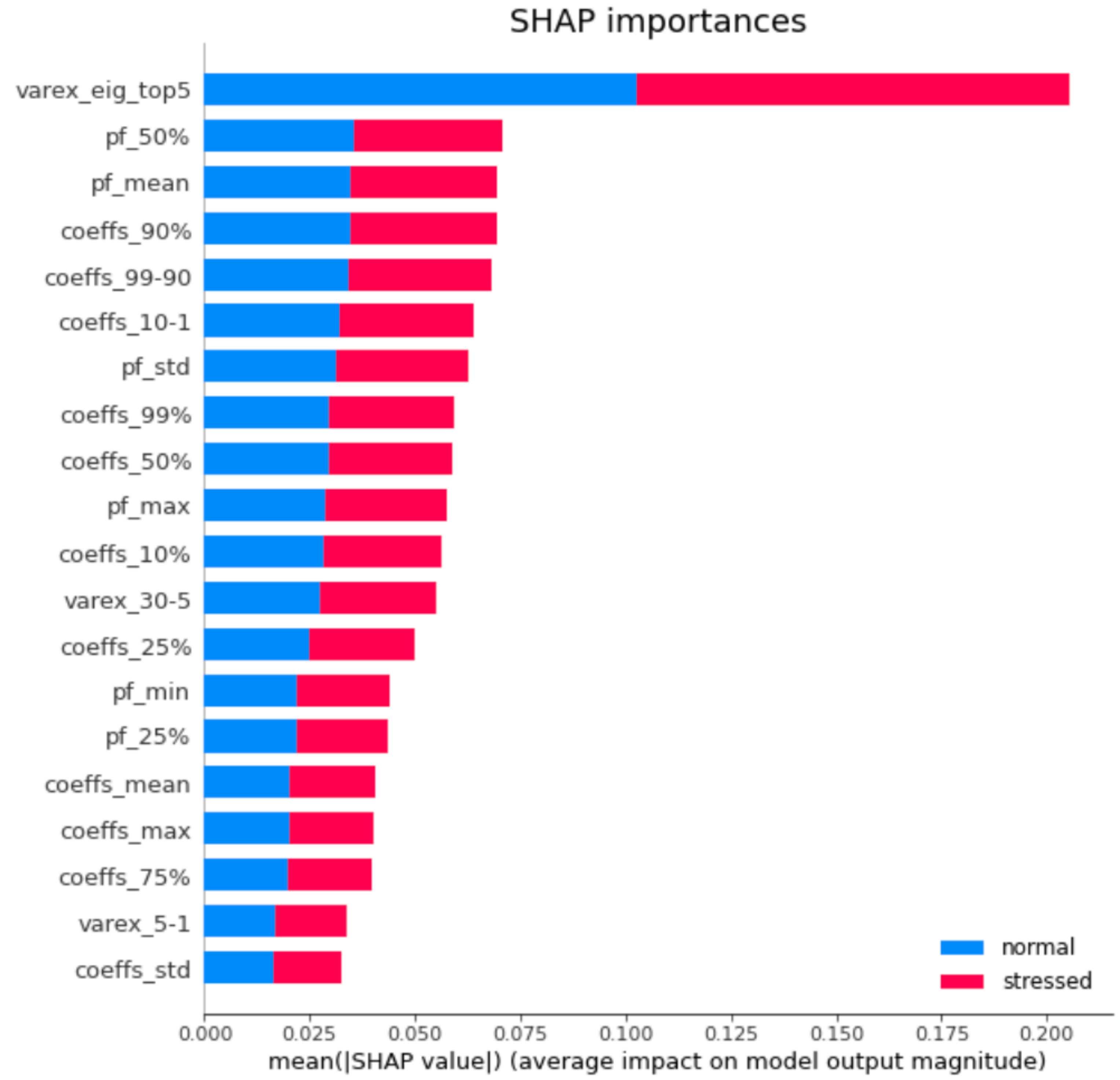
- Simple decision tree (**Random Forest Classifier**) model



GHENT UNIVERSITY

FACULTY OF ECONOMICS AND BUSINESS ADMINISTRATION

GAUGING THE MARKET REGIME USING CORRELATION MATRIX FEATURES: (4) INTERPRET THE MODEL



GAUGING THE MARKET REGIME USING CORRELATION MATRIX FEATURES: (5) BACKTESTING THE MODEL

- 100% SPDR Eurostoxx50 ETF Cumulative Returns
- **Can we time the market using our classifier?**

GAUGING THE MARKET REGIME USING CORRELATION MATRIX FEATURES: (5) BACKTESTING THE MODEL

```
dates = list(all_hist_corr.keys())

rets_taa = {}
all_predicted_regimes = {}
predicted_regime = ''

# TAA Strategy
for i, d in enumerate(dates):

    if d >= train_end_date:

        rets_taa[d] = (prices.loc[dates[i]] / prices.loc[dates[i-1]] - 1).values[0] if predicted_regime not in ['stressed']
        else 0.0
        predicted_regime = clf.predict(pd.DataFrame(compute_features_from_correl(all_hist_corr[d][0])).T)
        all_predicted_regimes[d] = predicted_regime

# 100% invested in ETF
rets_fully_invested = {}
for i, d in enumerate(dates):
        if d >= train_end_date:
                rets_fully_invested[d] = (prices.loc[dates[i]] / prices.loc[dates[i-1]] - 1).values[0]
```

GAUGING THE MARKET REGIME USING CORRELATION MATRIX FEATURES: (5) BACKTESTING THE MODEL

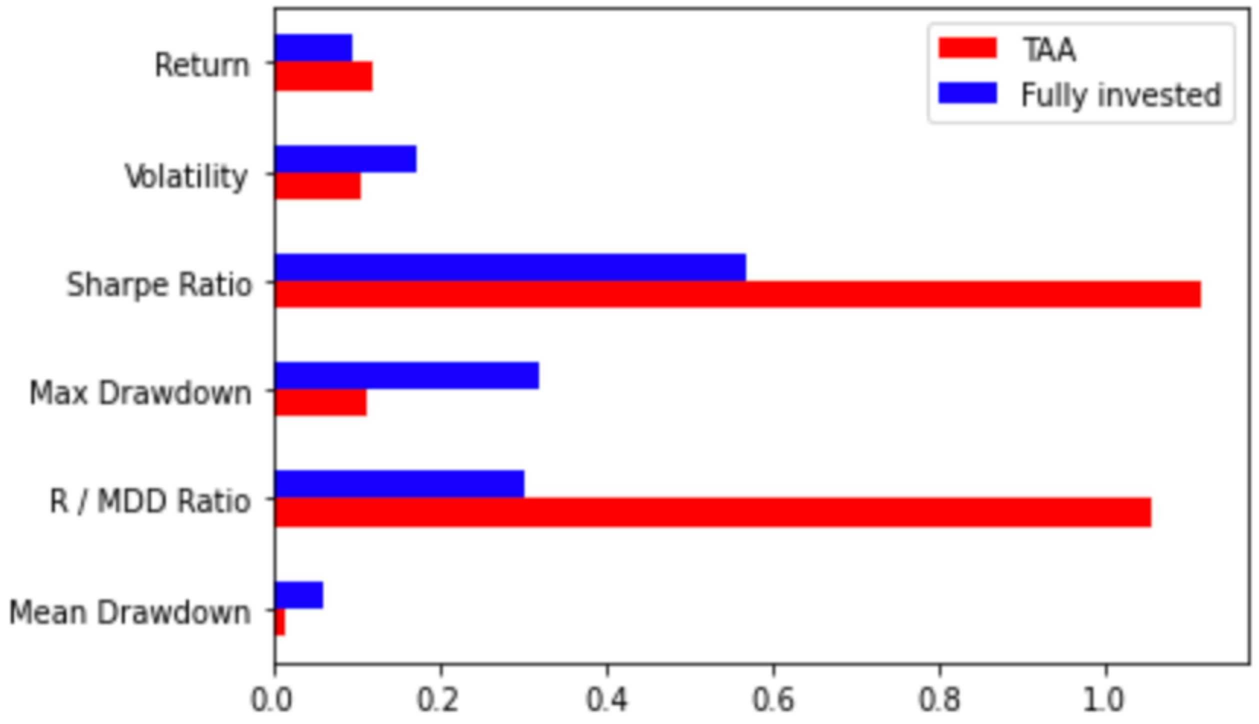
GAUGING THE MARKET REGIME USING CORRELATION MATRIX FEATURES: (6) EVALUATING THE MODEL

```python
def calc_stats(value):

    rets = value.pct_change()
    ret = rets.mean() * 12
    vol = rets.std() * np.sqrt(12)
    sharpe = ret / vol
    dd = value.cummax() - value
    max_dd = dd.max()
    return_over_maxdd = ret / max_dd
    mean_dd = dd.mean()
    pain = ret / mean_dd

    return {'Return': ret, 'Volatility': vol,
    'Sharpe Ratio': sharpe, 'Max Drawdown':
    max_dd, 'R / MDD Ratio':
    return_over_maxdd, 'Mean Drawdown':
    mean_dd, 'R / AvDD': pain}
```

	TAA	Fully invested
Return	0.118887	0.097246
Volatility	0.106479	0.171434
Sharpe Ratio	1.116535	0.567249
Max Drawdown	0.112514	0.319642
R / MDD Ratio	1.056644	0.304233
Mean Drawdown	0.016873	0.061382
R / AvDD	7.045932	1.584284



Emiel Lemahieu

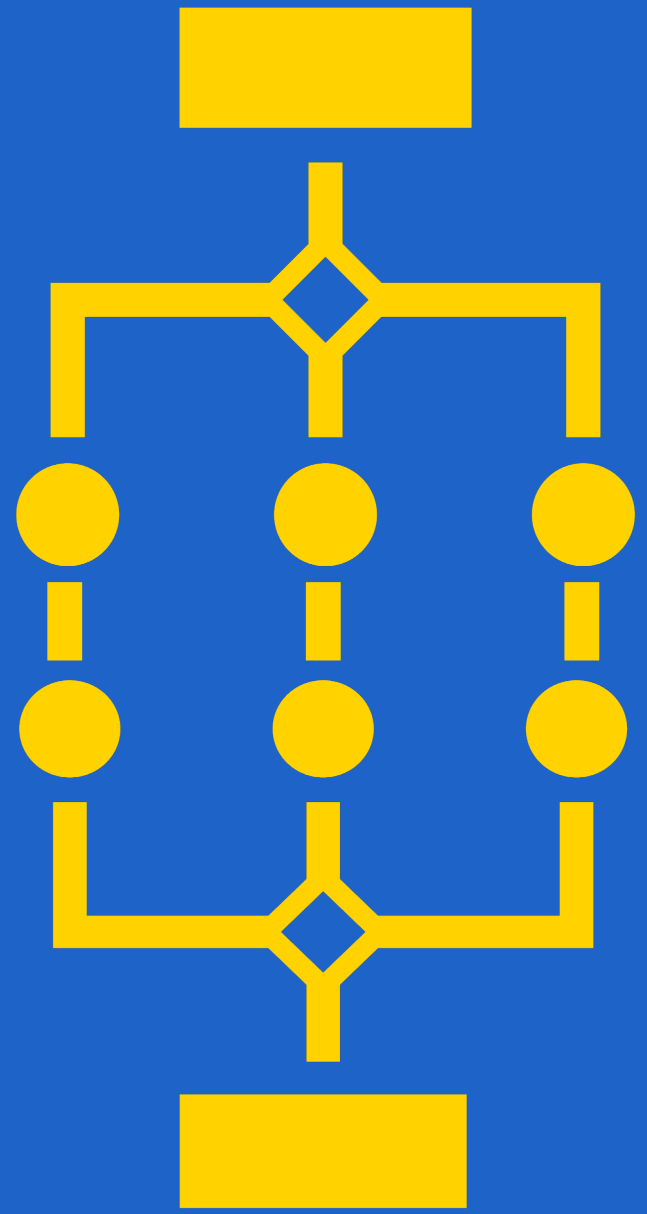
Quantitative Researcher

emiel.lemahieu@investsuite.com

www.linkedin.com/in/emiel-lemahieu

APPENDICES

A1: MACHINE LEARNING



GHENT UNIVERSITY

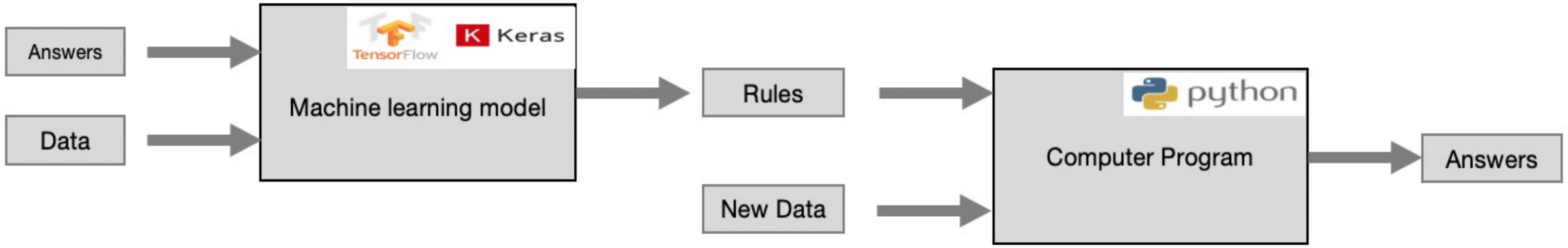
FACULTY OF ECONOMICS AND BUSINESS ADMINISTRATION

CORE IDEAS: LEARNING FROM DATA

- Learning from data

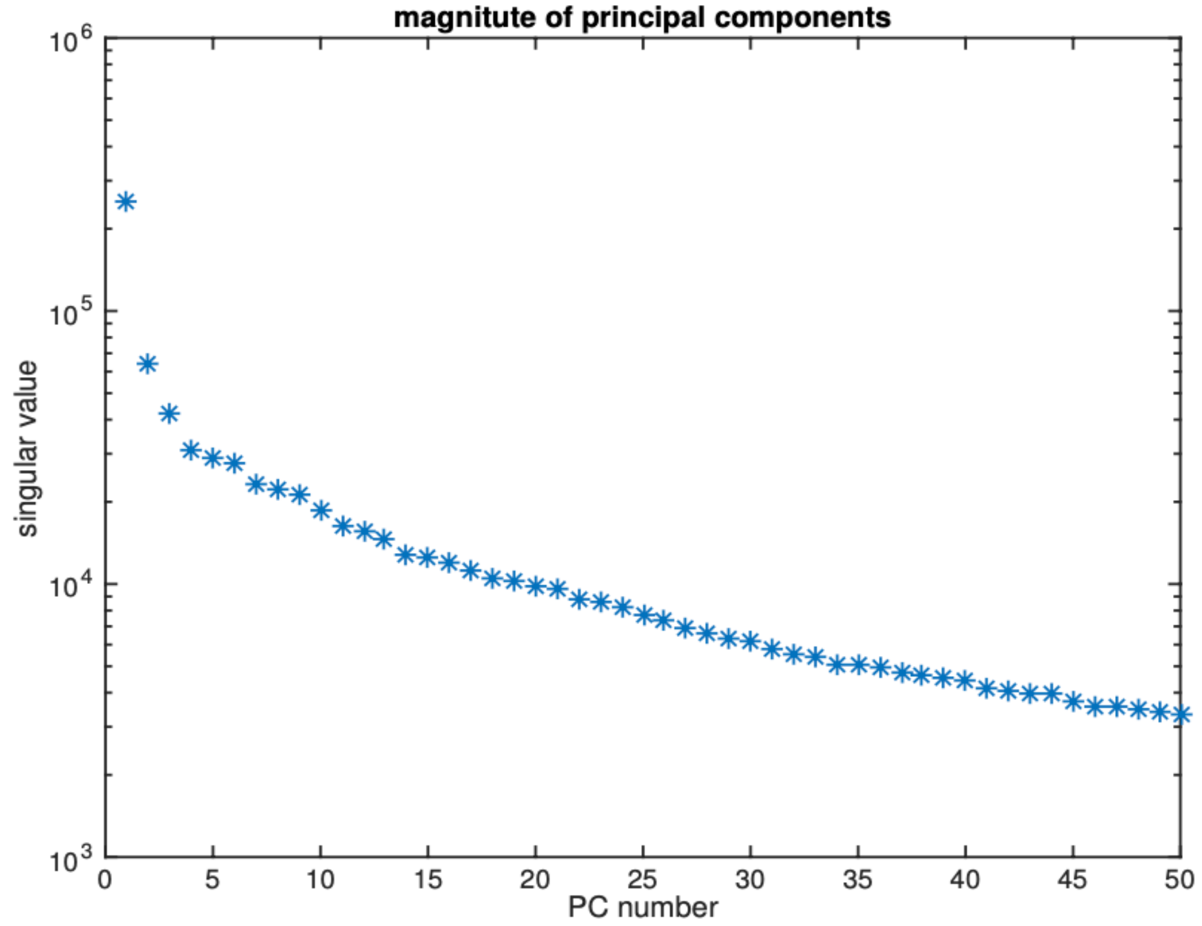
Model driven: we impose our view of how the data behaves in the form of rules

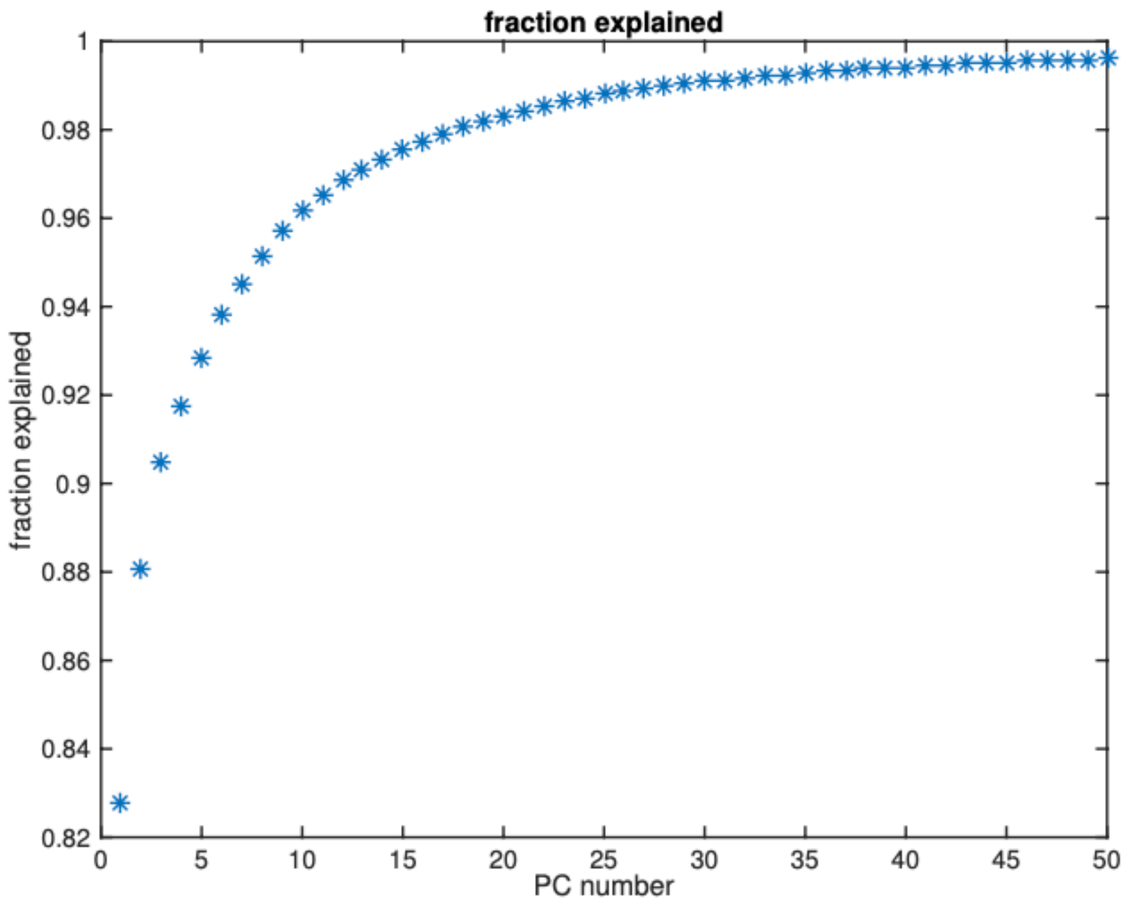
Data driven: the machine learns the data's view with rules as model output rather than model input



CORE IDEAS: APPROXIMATE SPARSITY

- High dimensional World, driven by low dimensional Rules
- ***Approximate Sparsity****, e.g. PCA**





**Formally: The sorted absolute values of the coefficients decay fast enough, i.e. the j^{th} largest coefficient (absolute value), $|\beta|_j \leq A j^{-a}, a \geq 1/2, \forall j$*

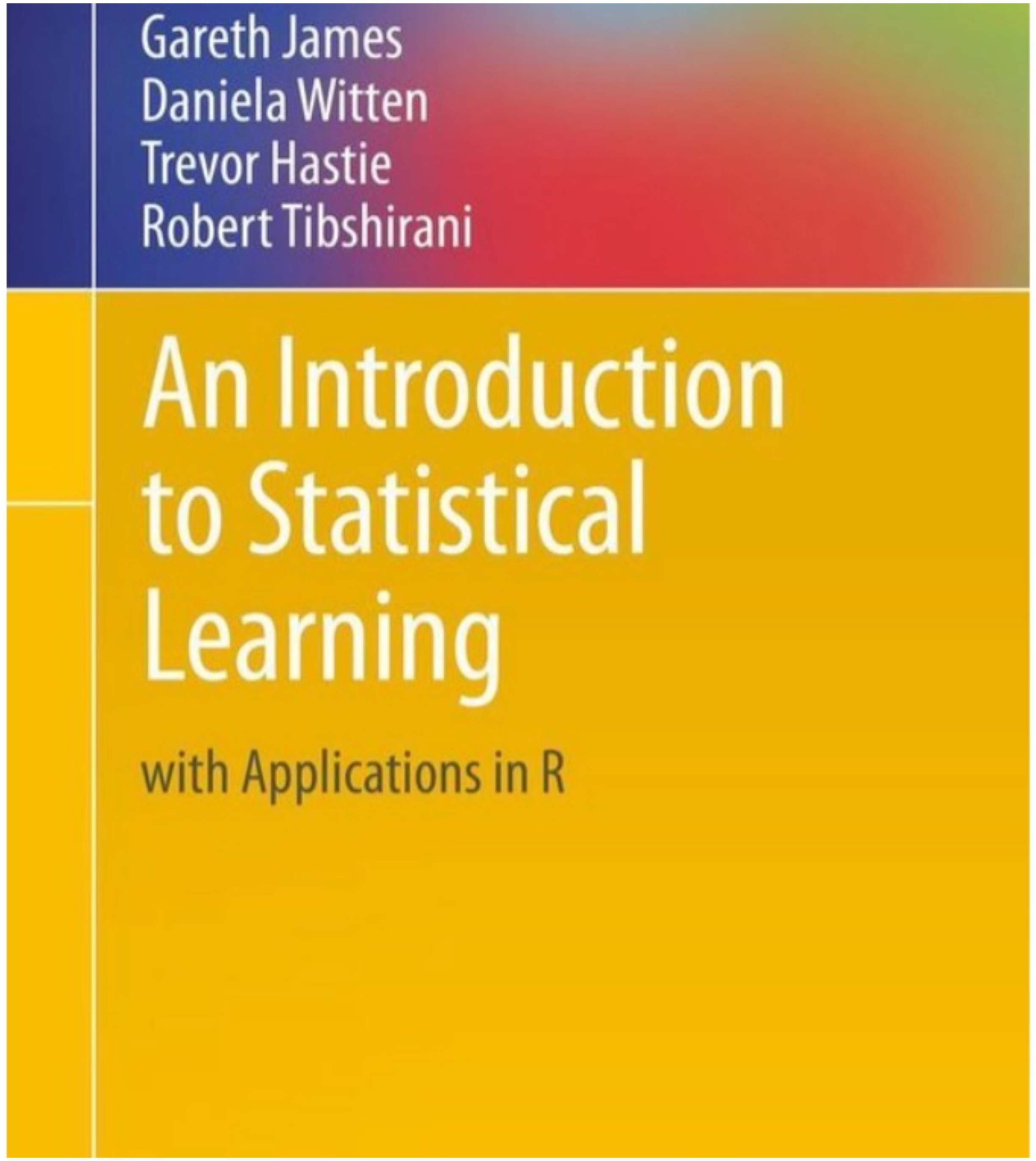
GHENT UNIVERSITY

FACULTY OF ECONOMICS AND BUSINESS ADMINISTRATION

TECHNIQUES 1/2 : RESOURCES

Welcome to the zoo!

- ***Generalized / penalized linear models***:
  - LASSO, Ridge, Elnet, …
- ***Tree-based methods***:
  - Decision Trees, Random Forests, Ensemble: Bagging, Boosting, …
- ***Graph-based methods***:
  - Graph theory (MST), …
  - Hierarchical Clustering methods, ...
- ***Neural Networks:***
  - Simple ANN or Feedforward NN (MLP – Multilayer Perceptron)
  - Convolutional networks (CNN, TCN)
  - Recurrent Networks (GRU, LSTM, WaveNet)
  - Graph nets (GNN, GraphSage)



Link to resource: An introduction to statistical learning

Original book: The elements of statistical learning: Data mining, inference, and prediction

CORE IDEAS: UNIVERSAL APPROXIMATION

ORIGINAL CONTRIBUTION

Approximation Capabilities of Multilayer Feedforward Networks

KURT HORNIK

Technische Universität Wien, Vienna, Austria

(*Received 30 January 1990; revised and accepted 25 October 1990*)

Abstract—*We show that standard multilayer feedforward networks with as few as a single hidden layer and arbitrary bounded and nonconstant activation function are universal approximators with respect to $L^p(\mu)$ performance criteria, for arbitrary finite input environment measures μ, provided only that sufficiently many hidden units are available. If the activation function is continuous, bounded and nonconstant, then continuous mappings can be learned uniformly over compact input sets. We also give very general conditions ensuring that networks with sufficiently smooth activation functions are capable of arbitrarily accurate approximation to a function and its derivatives.*

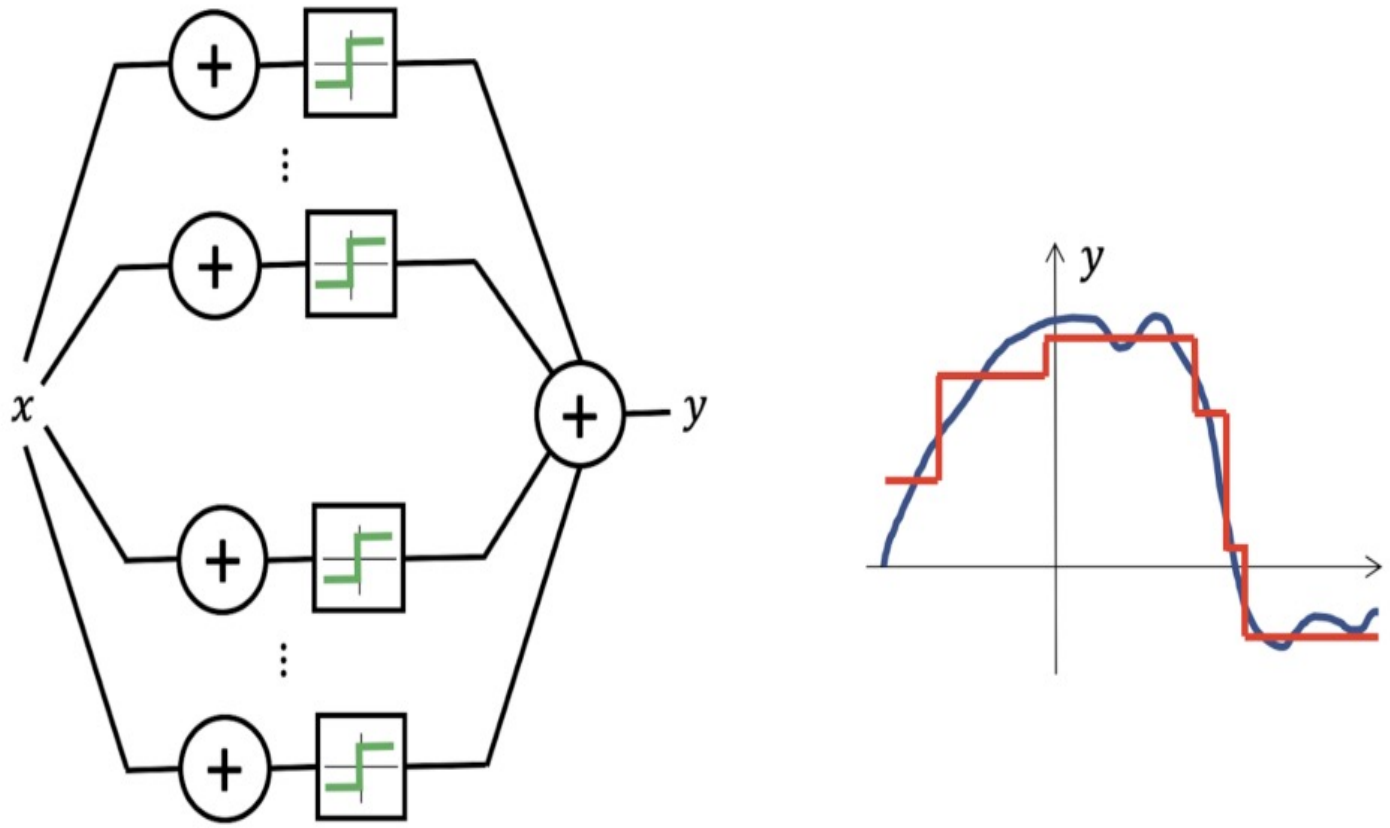


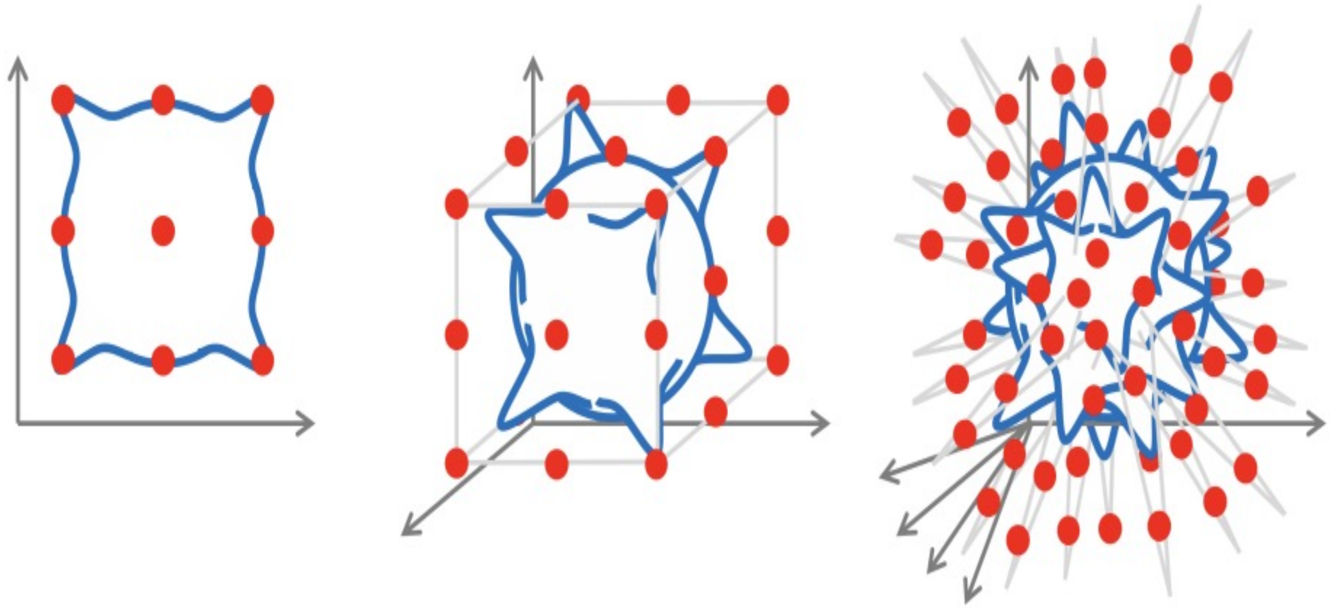
Figure 1: Multilayer Perceptrons (Rosenblatt, 1958), the simplest feed-forward neural networks, are universal approximators: with just one hidden layer, they can represent combinations of step functions, allowing to approximate any continuous function with arbitrary precision.

CORE IDEAS: CURSE OF DIMENSIONALITY

- Universal Approximation does **not** imply a prior-free world for the modeller due to scaling to higher dimensional approximating functions or manifolds

- Need to impose some form of ***regularity*** (= geometric priors like convolutions, cell structure or traditional dropout and shrunk coefficients to **reduce hypothesis space** to a smooth or less complex subset).

- But often far less stringent assumptions on the data before convergence is achieved
  -> main reason behind the succes of DL



GHENT UNIVERSITY

FACULTY OF ECONOMICS AND BUSINESS ADMINISTRATION

TECHNIQUES 2/2: RESEARCH

- Discriminative vs. ***generative*** modeling

- **Generative Adversarial Nets (GAN)**
  - Goodfellow et al 2014: Generative adversarial networks
  - In finance, Wiese et al. 2020: Quant GAN
- **Restricted Boltzmann Machines (RBM)**
  - First introduced as Harmoniums by Smolensky 1986: Harmonium
  - Used as market generators in Kondratyev and Schwarz (2019): The Market Generator
- **Variational Autoencoders (VAE)**
  - Welling et al. 2013: Auto-encoding variational bayes
  - Applied to sampling market paths: Buehler et al. 2020 A data-driven market simulator for small data environments

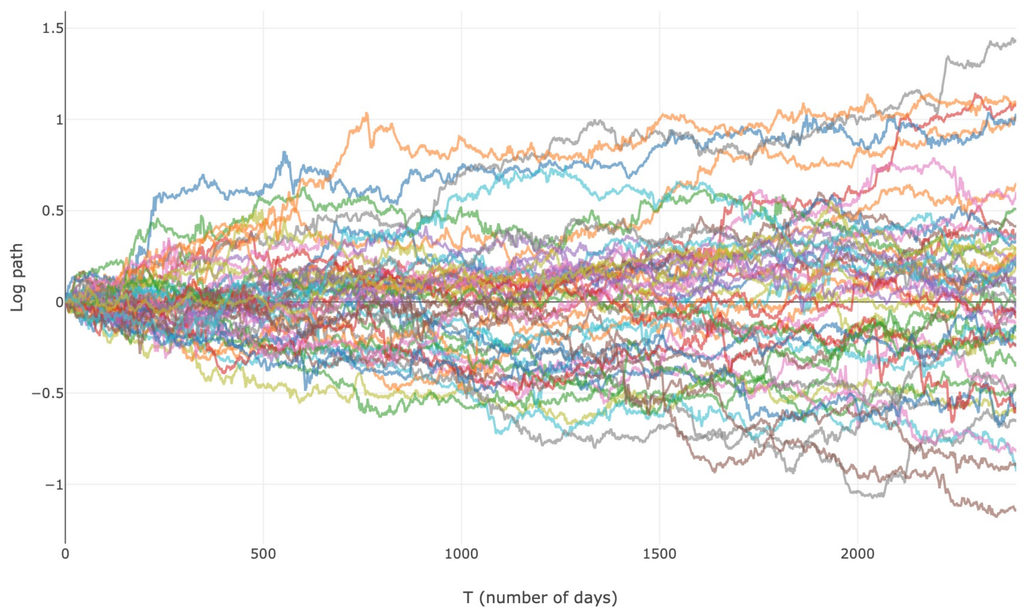


Figure A.10: 50 generated log paths

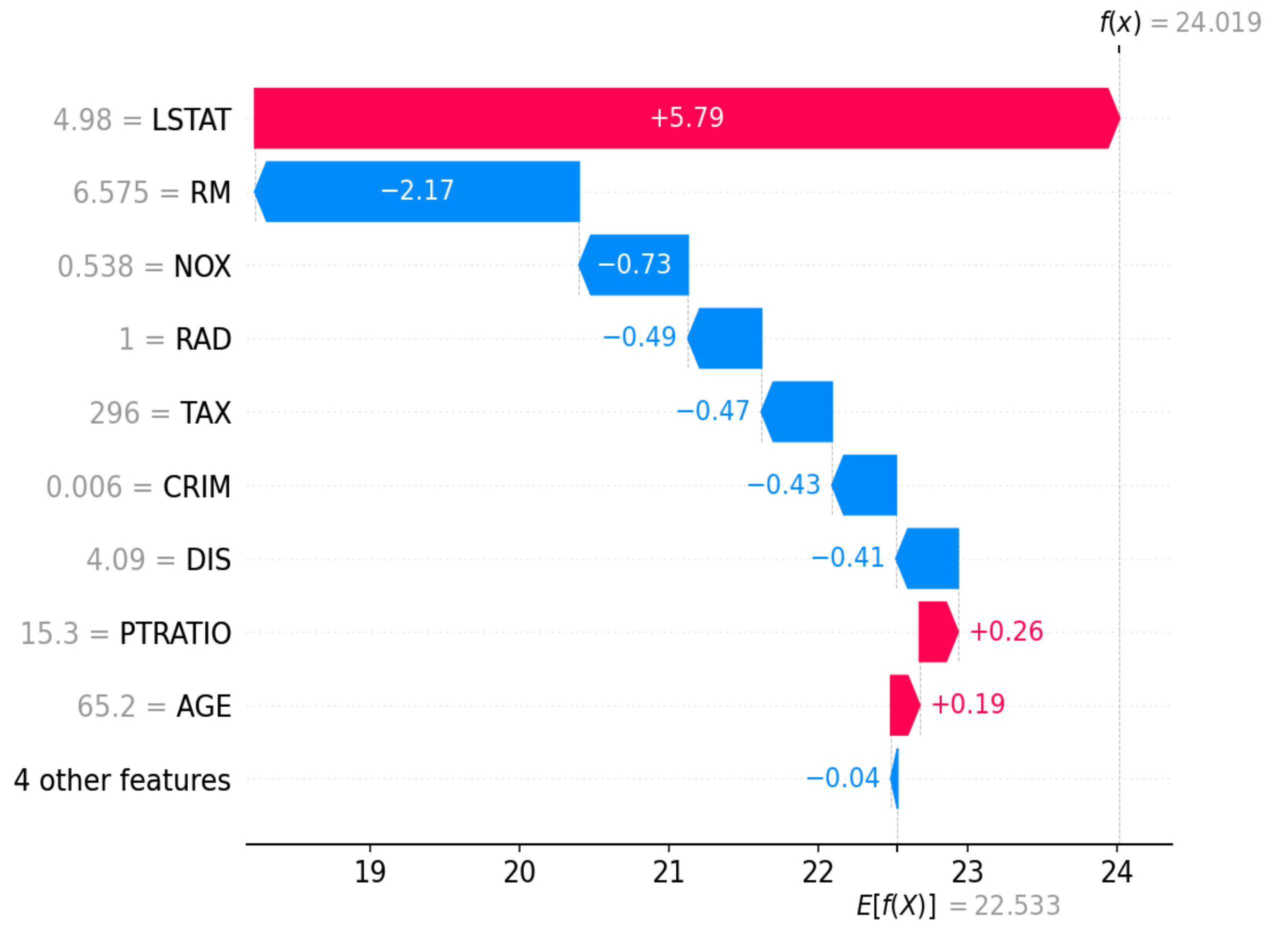
SHAPLEY VALUES

https://github.com/slundberg/shap

- **SHAP (SHapley Additive exPlanations)** is a game theoretic approach to **explain the output of any machine learning model**.

- Easy-to-use package with plug-and-play code for tree-based (ensemble) methods (XGBoost/LightGBM/CatBoost/scikit-learn/pyspark models). We used the scikit-learn ('sklearn') **RandomForrestClassifier**.

- Also easy examples for **Natural language** (e.g. transformers) and **Deep learning** (TensorFlow/Keras models) models!



Example: SHAP values

GHENT UNIVERSITY

FACULTY OF ECONOMICS AND BUSINESS ADMINISTRATION

SHAPLEY VALUES BEESWARM: FEATURE CONTRIBUTION TO PREDICTION, VISUALIZED PER OBSERVATION