AXYON.AI

AI & deep learning solutions
for asset management

VaR Estimation with conditional GANs and GCNs

IS4: NEW TRENDS IN FINANCE INDUSTRY

SIMAI 2020+2021
Parma, 30 Aug - 3 Sep 2021

Agenda

1. Introduction [5’]
    - Company
    - Research
2. VaR estimation with conditional GANs and GCNs [30’]
    - Project goals
    - Methods
    - Results
    - Takeaways
    - Further research
3. Q&A [10’]

Introduction

Company

Axyon AI is an Italian fintech company on a mission to
build a factory of AI solutions for investment management

4	80%	10+	>150 k
Years old	Tech employees	Research projects	GPU hours per year

Partners

UNIMORE
UNIVERSITÀ DEGLI STUDI DI MODENA E REGGIO EMILIA

NVIDIA

IBM

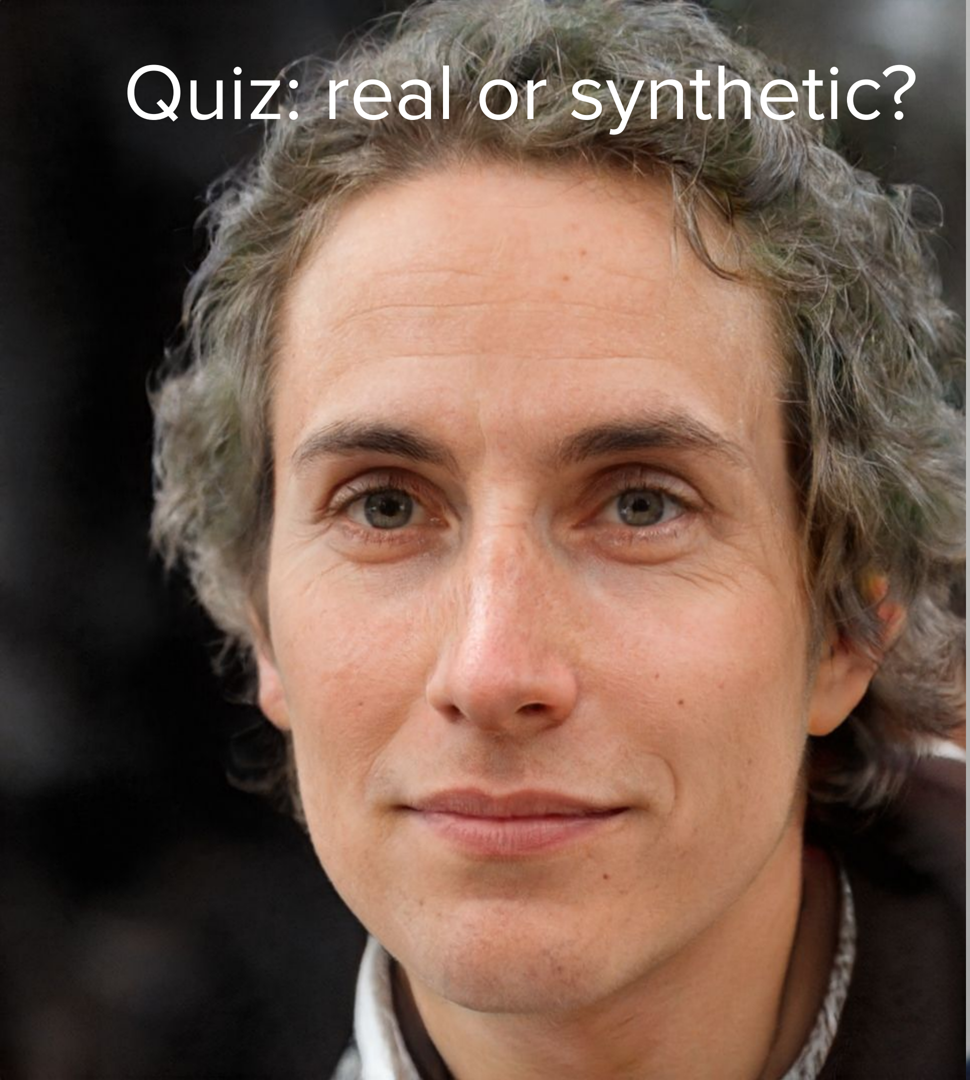
Microsoft

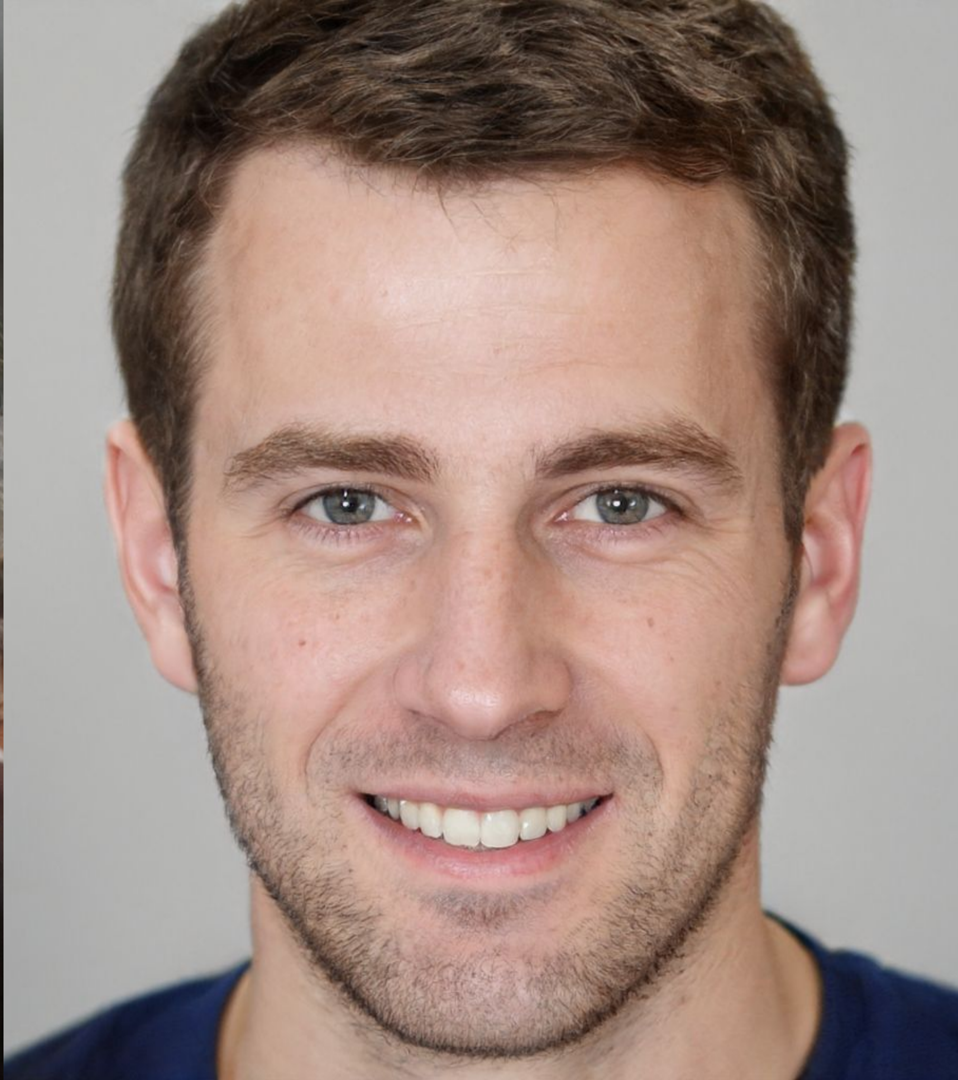
Investors

UniCredit

ING

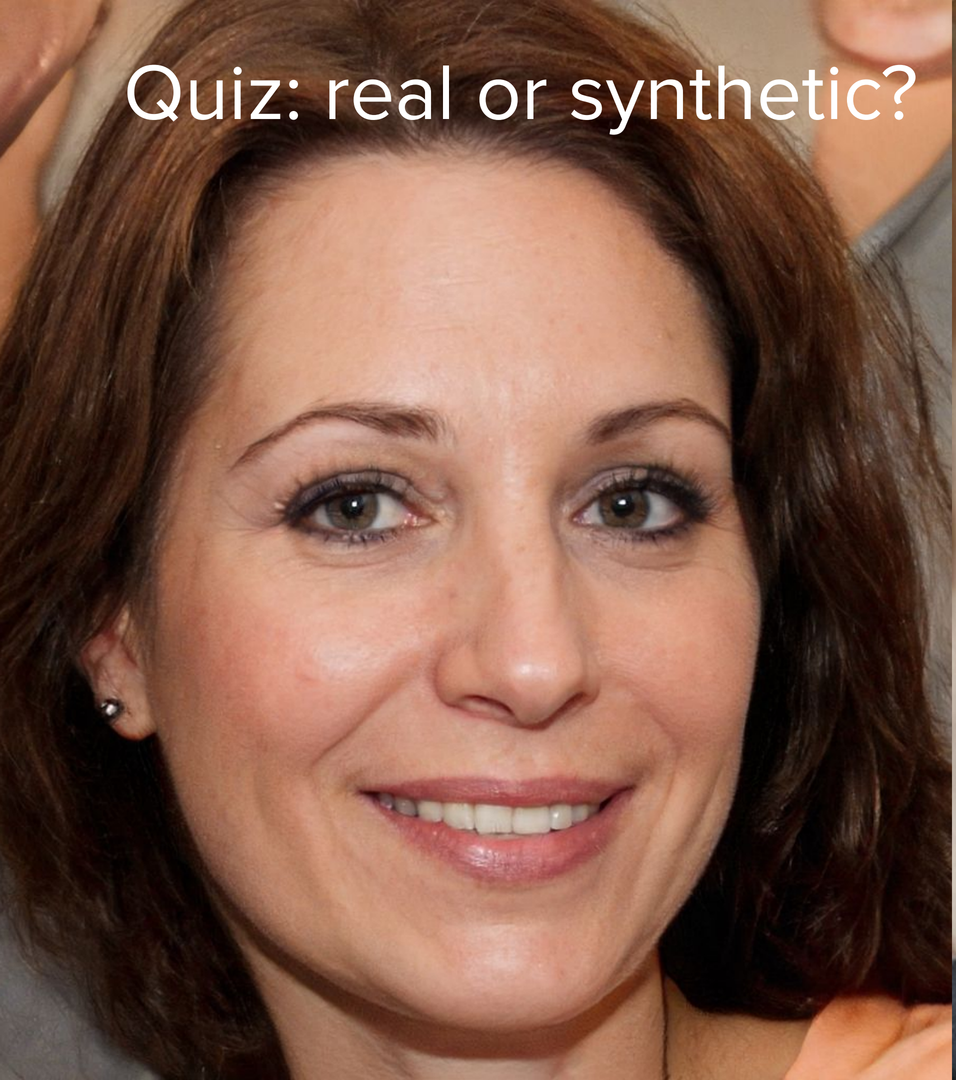
Quiz: real or synthetic?





Quiz: real or synthetic?

Quiz: real or synthetic?



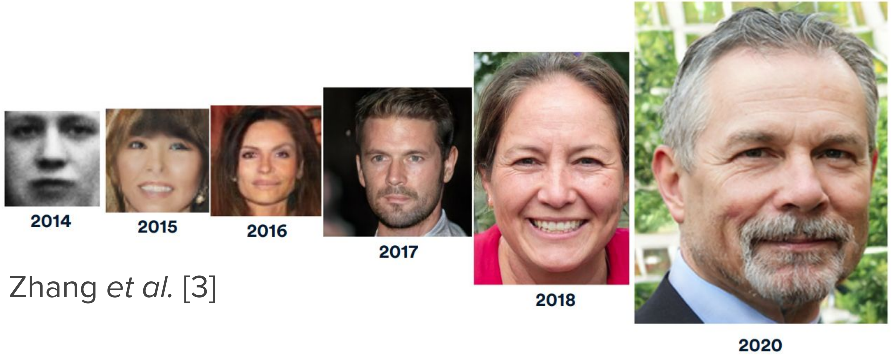
Quiz: real or synthetic?

Solution

All images in the previous slides are synthetically generated!

- Don't believe me, try for yourself at https://www.thispersondoesnotexist.com/ (Karras *et al.* [1])
- Seminal paper: Goodfellow *et al.* "Generative Adversarial nets" [2]

> *[GANs are] the most interesting idea in the last 10 years in Machine Learning, in my opinion.*
> Yann LeCun, 2018 ACM Turing Award Laureate, Chief AI Scientist, Facebook - 28 July 2016



Zhang *et al.* [3]

AXYON.AI

Project Goals and Scope

Why is generative modeling interesting?

We only know the realized history of financial time series. What if we could **generate alternative realistic paths** or just **model the conditional distribution of returns without making any assumption**?

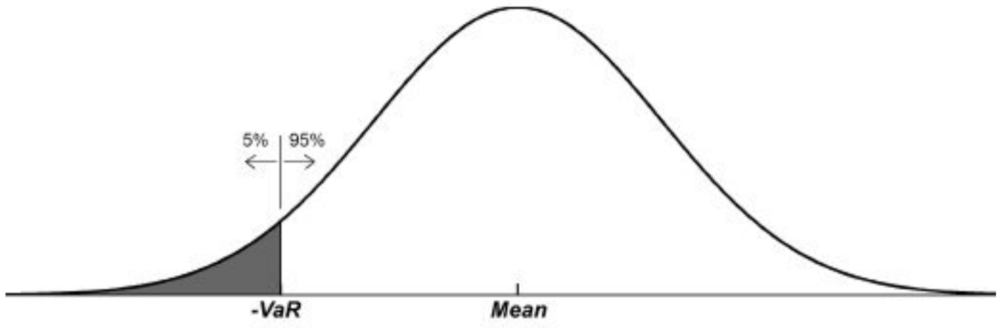
Potential Applications

- Data Augmentation
- Strategy Robustness Tests and Scenario Simulations
- Derivatives Pricing
- Outliers Detection
- **Risk Management (Value At Risk estimation)**
- ...

AXYON.AI

Project Scope and Value at Risk

- With the aim to start with a *relatively* simple challenge we decided to **model the distribution of returns with Generative Adversarial Networks (GANs)** in order to estimate the Value at Risk (VaR) of a portfolio.
- To show that GANs can potentially become a promising technique for risk management, we decided to **compare our approach with a traditional GARCH baseline**.
- What is VaR?



AXYON.AI

A note about our baseline: the GARCH

- If we assume that the **conditional variance of a time series of stock returns is specified as linear function of both past squared returns and past conditional variances** we say **we are using a generalized autoregressive conditional heteroskedasticity model (GARCH)**.

$$\sigma^2_{t|t-1} = \alpha_0 + \sum_{i=1}^{q} \alpha_i r^2_{t-i} + \sum_{j=1}^{p} \beta_j \sigma^2_{t-j|t-j-1}$$

- We basically estimate the variance at time t using a *trained GARCH model* and assume this is the variance of the chosen distribution we will use to compute the VaR.
- Why do people use this approach? Well, it's **fast and easy** to interpret and implement BUT we **must rely on our assumptions on model specification and distribution of returns, which may be wrong**.

AXYON.AI

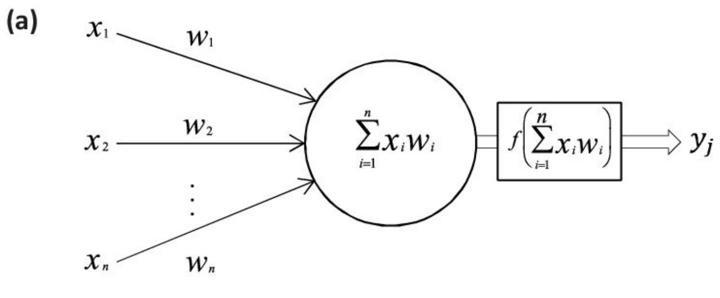
VaR estimation with conditional GANs and GCNs

AXYON.AI

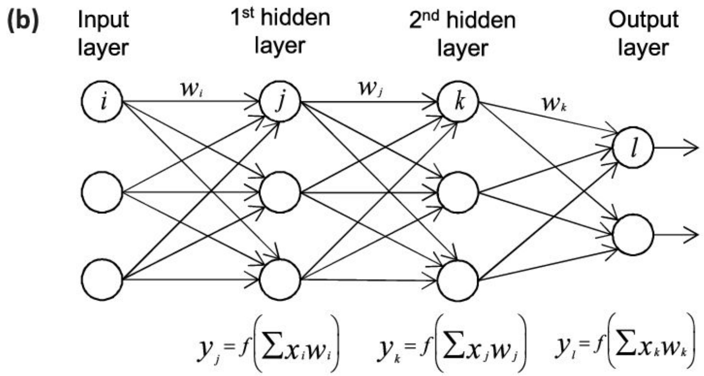
Artificial Neural Networks (ANNs)

A brief introduction

ANNs are mathematical models loosely inspired by the structure and operation of biological learning systems.



A single perceptron



A fully connected NN built of a variable number of perceptrons, organized in layers

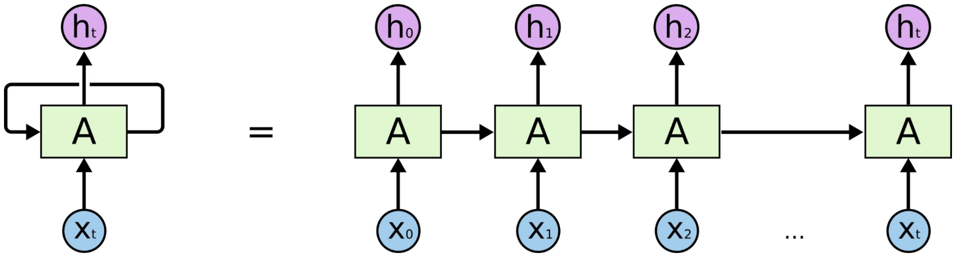
AXYON.AI

Long Short Term Memory (LSTM)

Recurrent Neural Networks

RNNs are networks with loops in them, allowing information to persist.

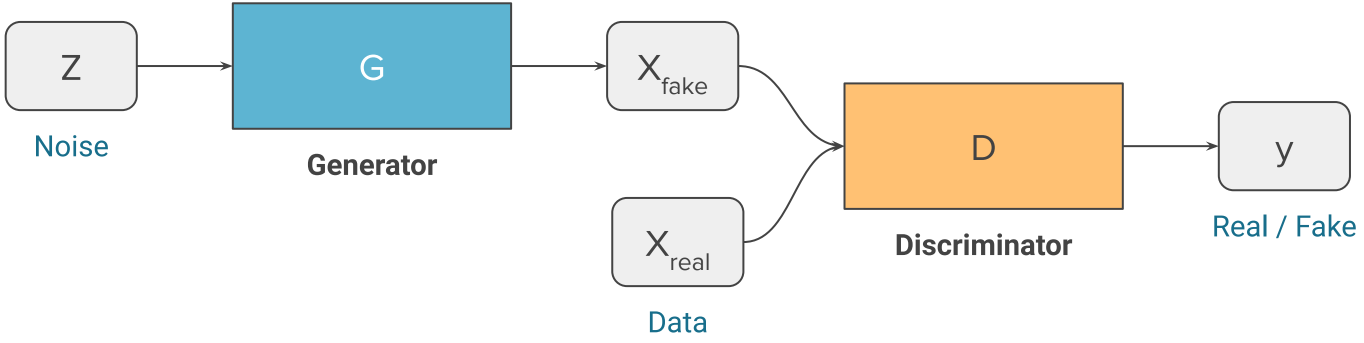
LSTMs are a special kind of RNN, capable of learning long-term dependencies.



The key to LSTMs is the cell state, the horizontal line running through the top of the diagram. The LSTM does have the ability to remove or add information to the cell state, carefully regulated by structures called gates.

Generative Adversarial Networks

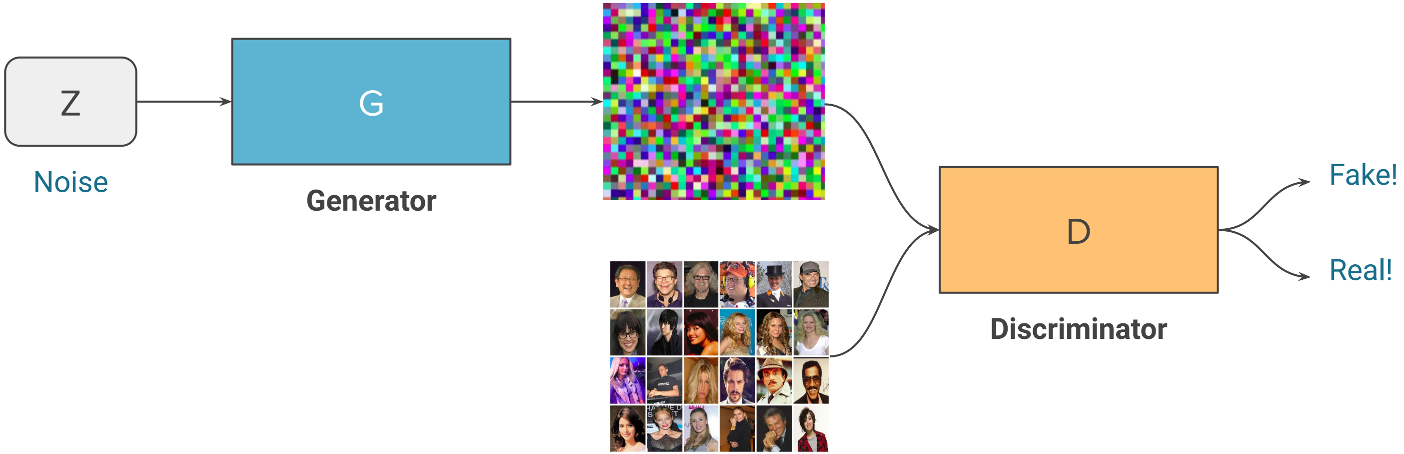
Base structure



AXYON.AI

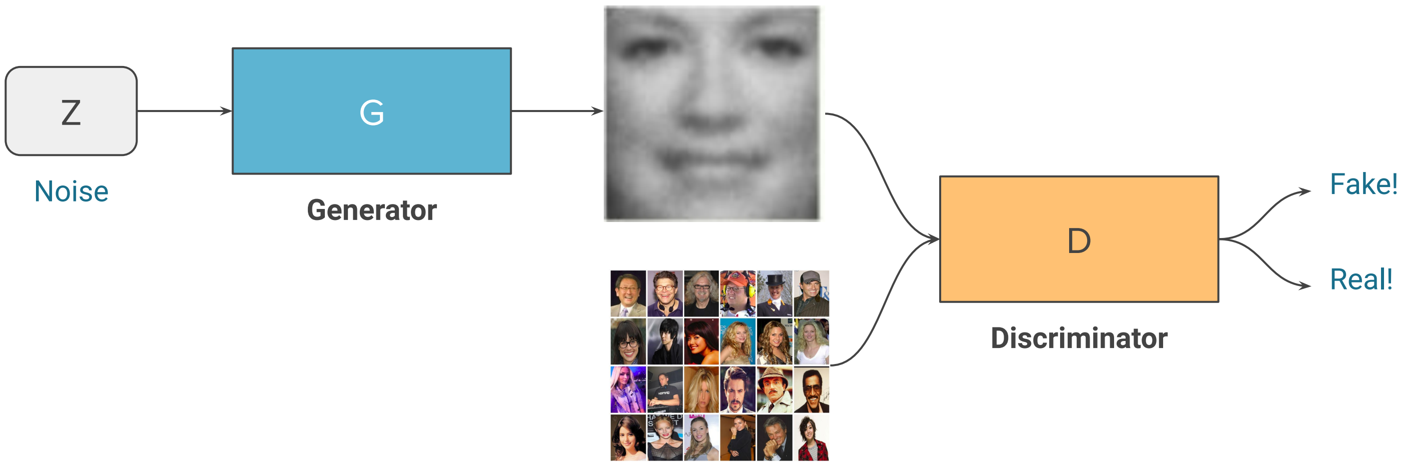
Generative Adversarial Networks

Example: image generation



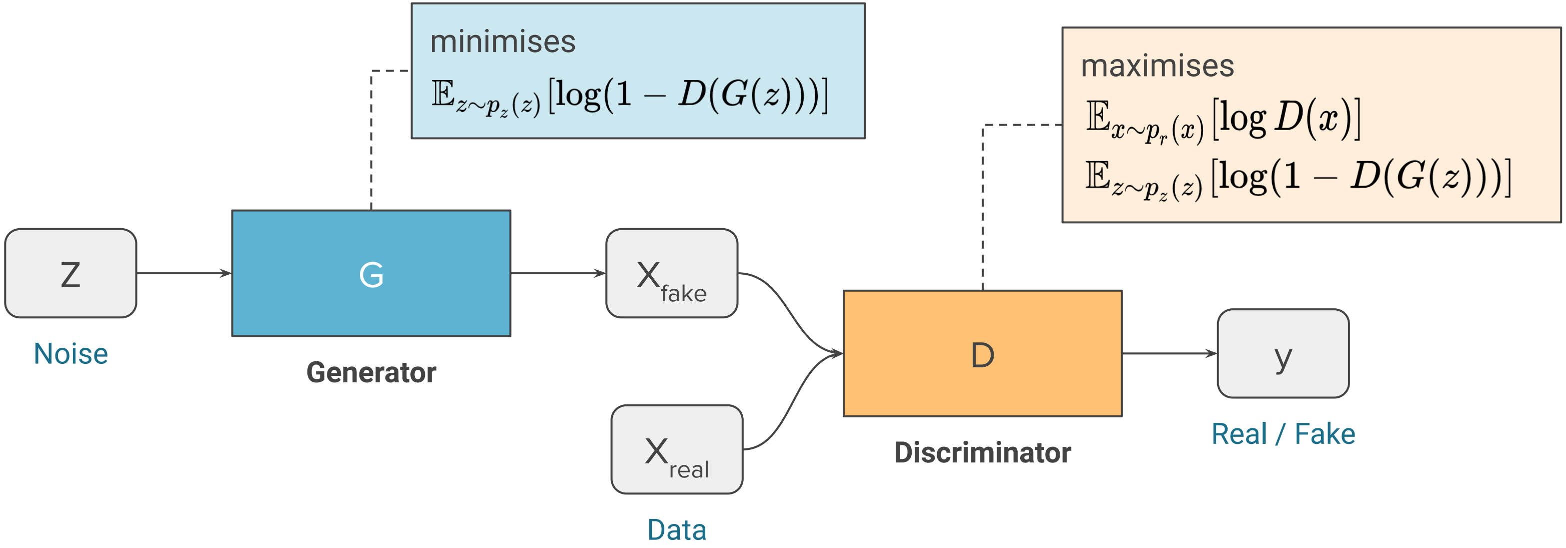
Generative Adversarial Networks

Example: image generation



Generative Adversarial Networks

Mathematical framework

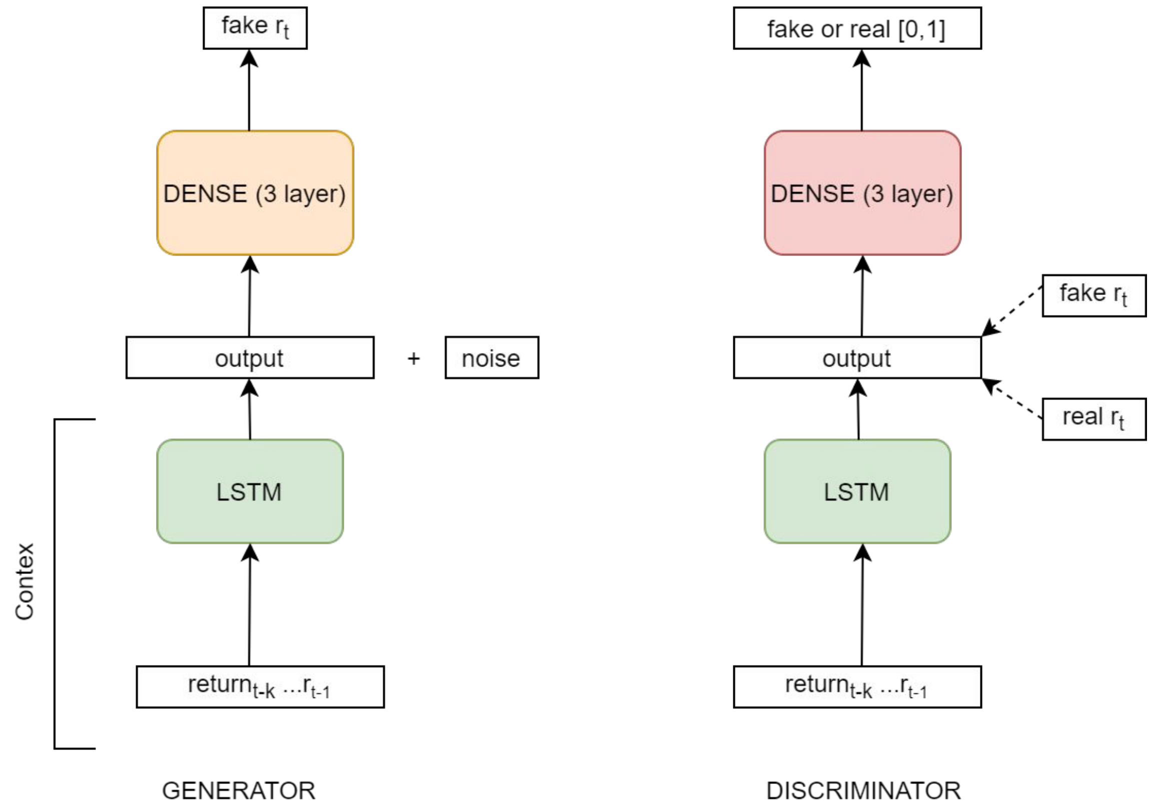


It's a min-max game!

$$\min_G \max_D L(D, G) = \mathbb{E}_{x \sim p_r(x)}[\log D(x)] + \mathbb{E}_{z \sim p_z(z)}[\log(1 - D(G(z)))]$$
$$= \mathbb{E}_{x \sim p_r(x)}[\log D(x)] + \mathbb{E}_{x \sim p_g(x)}[\log(1 - D(x)]$$

⇒ D learns p (y | X), G learns p(X)

Univariate CGAN model

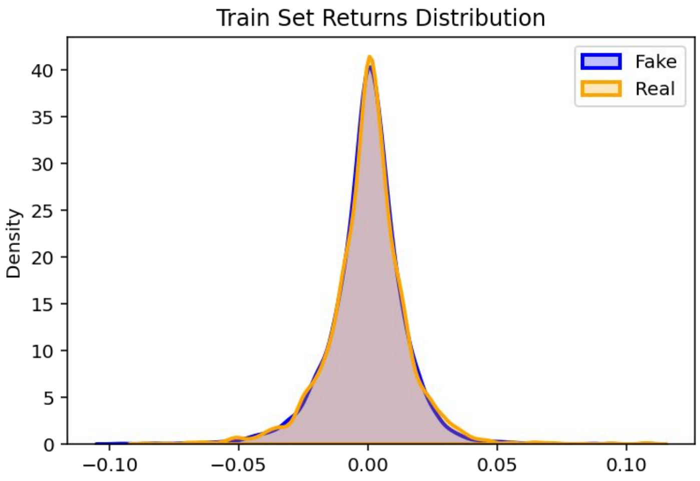


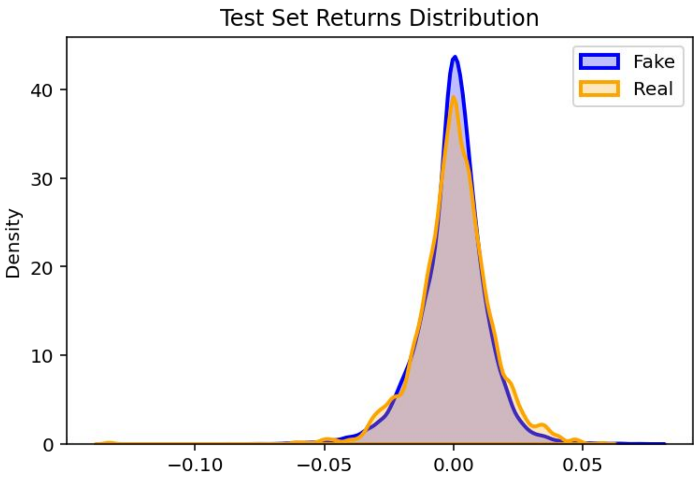
AXYON.AI

Results

Return distributions

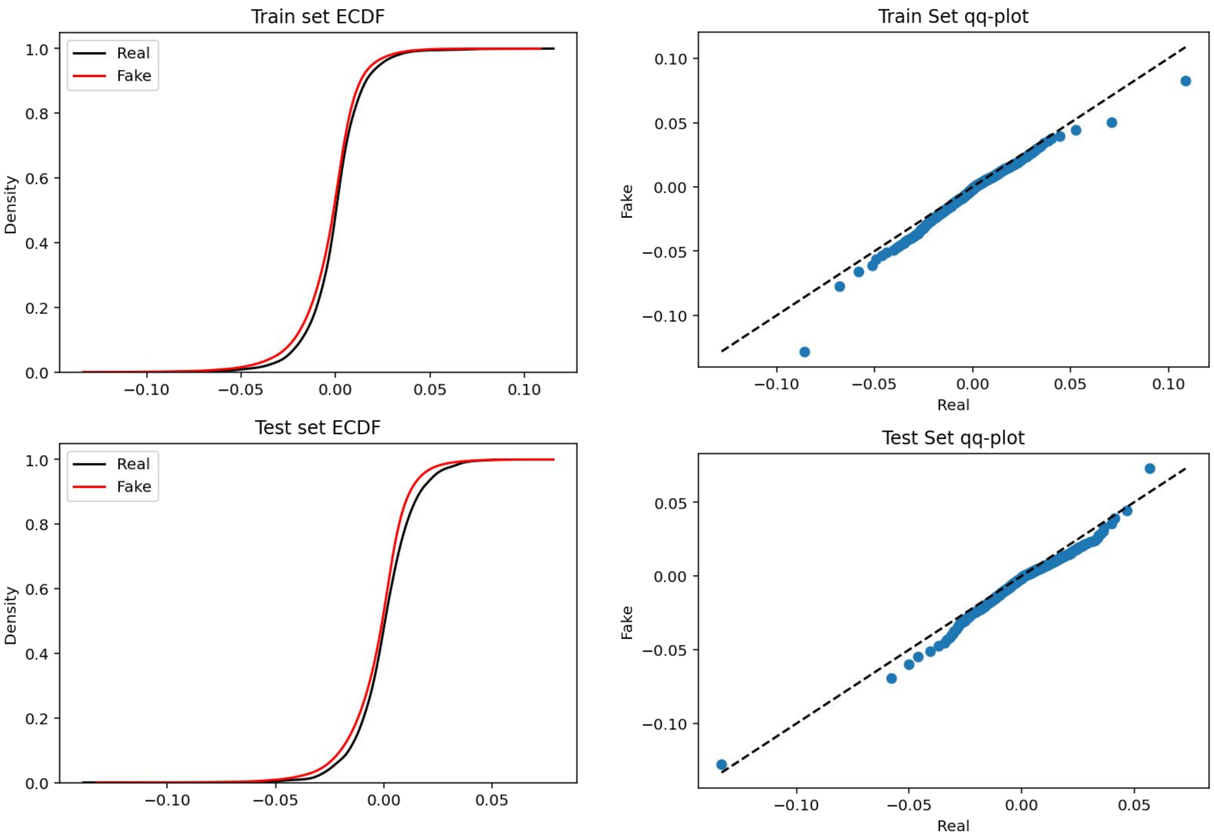
The model is trained on daily FTSE MIB returns from 2000 to 2012 for 3000 epochs and tested on the following years (2013-2018).





AXYON.AI

Return distributions



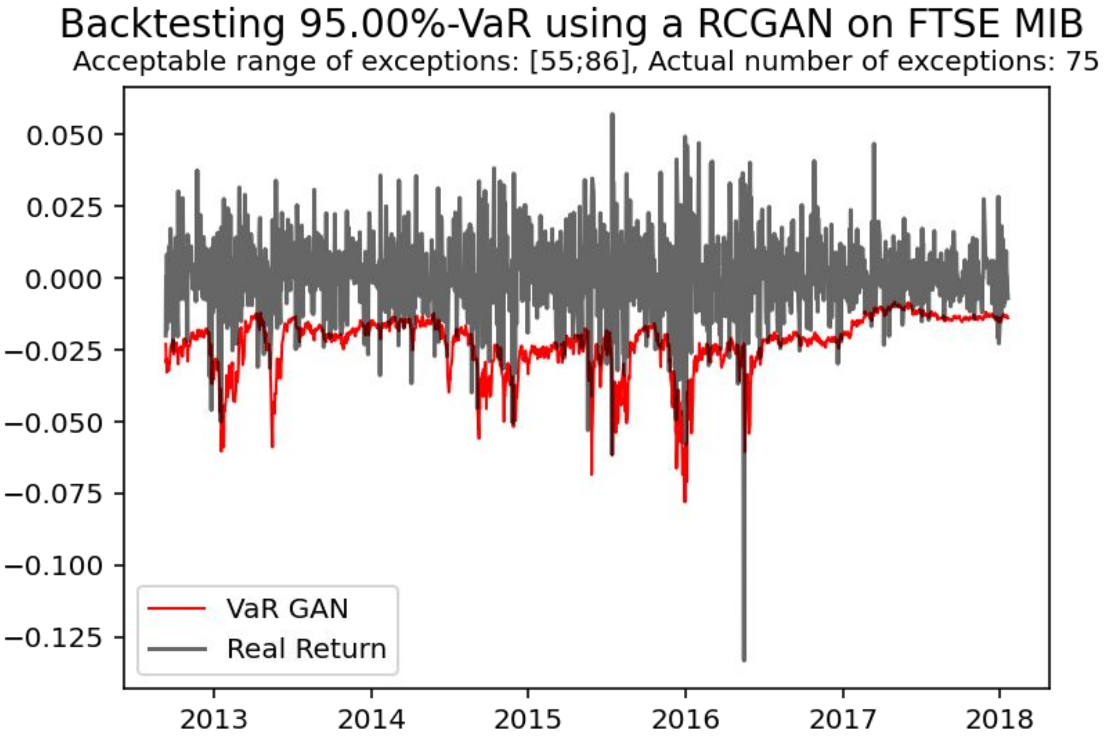
Backtesting procedure

1. **Train** the Generator and Discriminator for 3000 epochs
2. **Test** on the test set according to the below procedure [4]

```
for each day t in test set:
    while t in current month:
        generate 3200 samples for t+1
        compute VaR
    update the Generator and Discriminator on the last 5 months for 2 epochs
```

Backtesting results

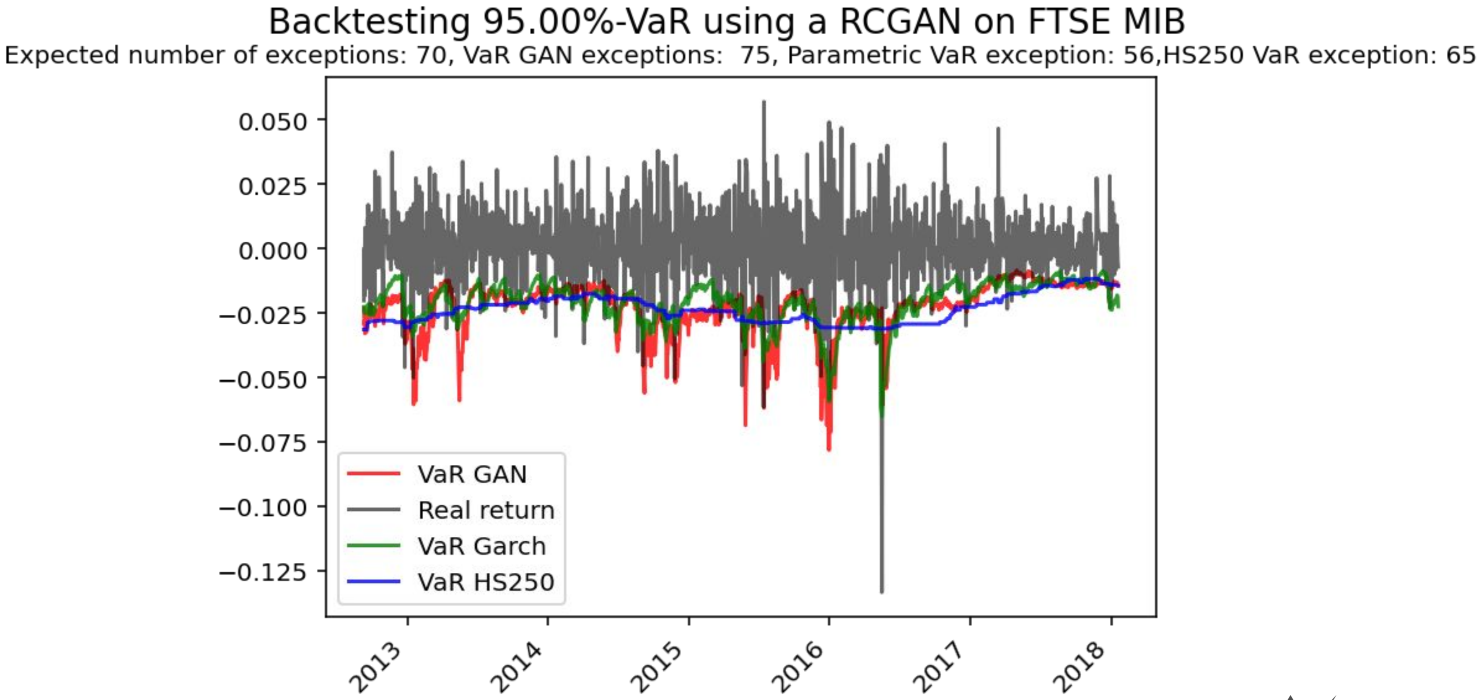
Univariate RCGAN



AXYON.AI

GAN vs GARCH

The two methods have similar VaR estimation.



AXYON.AI

Model validation

We use three different tests for our model validation [5]:

- The *Proportion of Failure* (POF) Test examines how many times a VaR is violated over a given span of time.
- The *Christoffersen's Interval Forecast* Test measures whether the probability of observing an exception on a particular day depends on whether an exception occurred on the previous day.
- The *Time Between Failures* test incorporates the time information between all the exceptions in the sample, the number of periods between exceptions should be independent.

		p-value		
$Test$	H_0	$GARCH(1,1)-norm$	$CGAN$	$HS250$
POF	$Pr[I_t(\alpha)=1]=\mathbb{E}[I_t(\alpha)]=\alpha$	0.075	0.544	0.535
CCI	$\mathbb{E}[I_t(\alpha)\|I_{t-1}(\alpha)]=\alpha$	0.030	0.991	0.037
TBFI	$\mathbb{E}[I_t(\alpha)\|\Omega_{t-1}]=\alpha$	0.702	0.01	0.0

Is it worth it?

Advantages

+ GAN can be viewed as universal approximators of probability distributions, useful when the underlying distribution is very complex (i.e financial time series)
+ No assumption needs to be made
+ Robust results for unconditional coverage and independence (though not for TBF)

Disadvantages

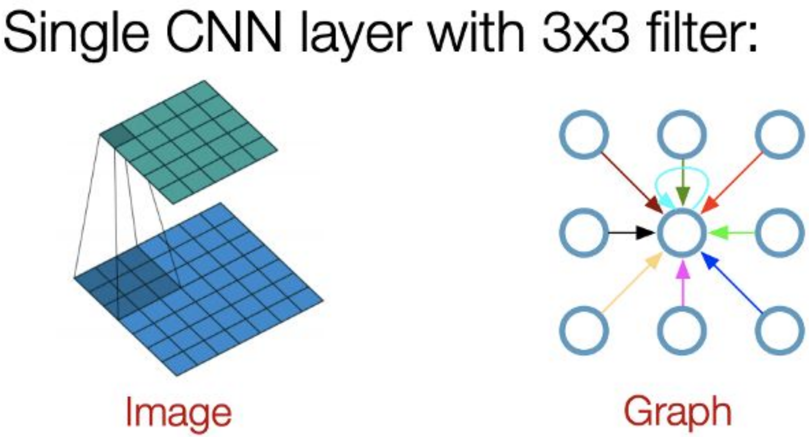
- Difficult to find a good set of hyperparameters and to tune the model
- Long training, other methods are faster
- Stochastic training requires careful setup for reproducibility
- As GARCH, the model is univariate…
  …or is it?

Multivariate RCGAN

AXYON.AI

Graph Convolutional Network

GCNs generalize classical Convolutional Neural Networks (CNN) to the case of graph-structured data, images have a fixed structure, Graphs are much more complex. [6]



The graph can be represented by the **adjacency matrix** A, it is a square matrix whose elements indicate whether pairs of vertices are adjacent, i.e. connected, or not.
In the simplest case, A_{ij} is 1 if there is a connection from node i to j, and 0 otherwise.

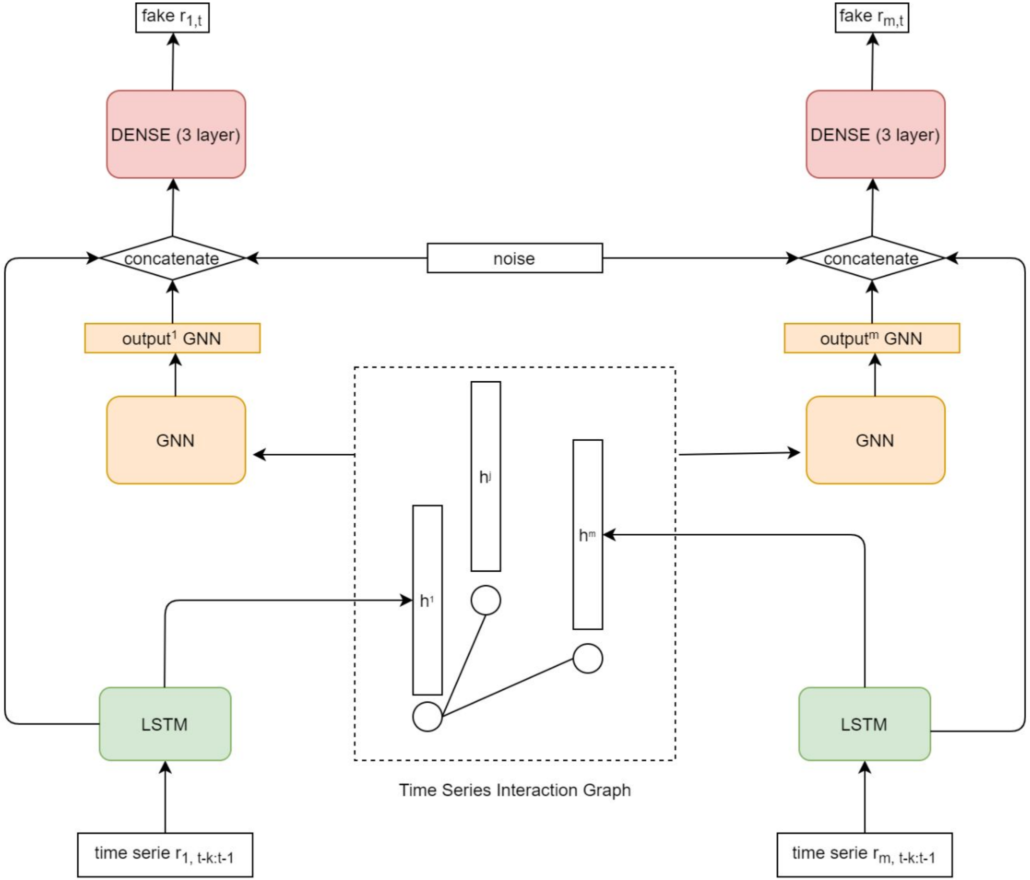
AXYON.AI

Graph Convolutional Network

In mathematical terms:

$$H^{(l+1)} = \sigma\left(\hat{D}^{-1/2}\hat{A}\hat{D}^{-1/2}H^{(l)}W^{(l)}\right)$$

where $W^{(l)}$ is the weight parameters with which we transform the input features into messages ($H^{(l)}W^{(l)}$). Finally, to take the average instead of summing, we calculate the matrix D which is a diagonal matrix with D_{ii} denoting the number of neighbors node i has. σ represents an arbitrary activation function



Time Series Interaction Graph

The LSTM nets capture the comprehensive influence of the temporal dependencies within each time series, they extract **temporal dependencies**.

The GCN (graph convolutional network) capture the interactional dependencies between time series, they extract **interactional dependencies**. [7]

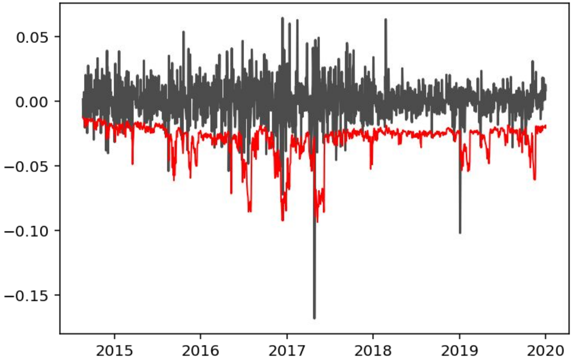
AXYON.AI

Results for 5 stocks (from EUROSTOXX50)

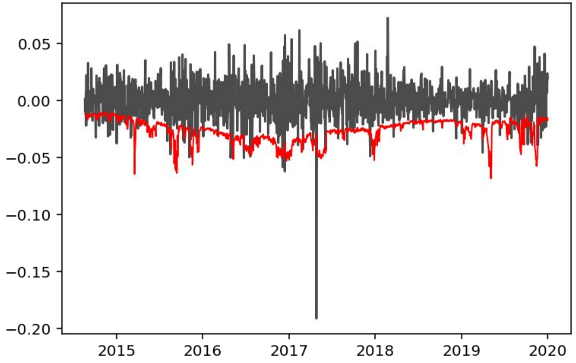
$$\begin{bmatrix} 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 \end{bmatrix}$$

Adjacency matrix, 1 if correlation between the two assets larger than threshold (corr > 0.7)

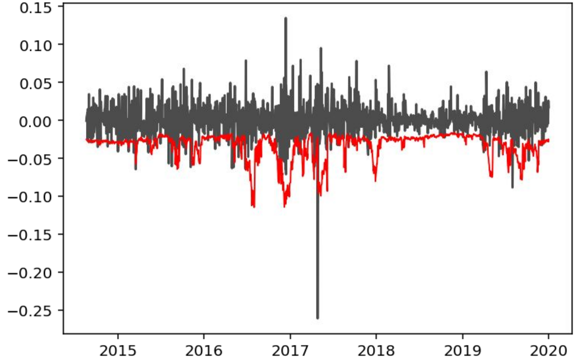
F:MIDI - Actual number of exceptions: 51



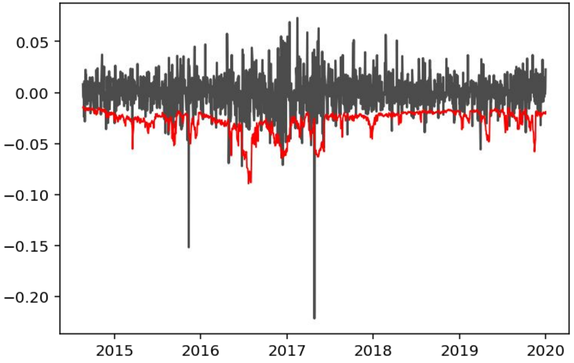
F:BNP - Actual number of exceptions: 85



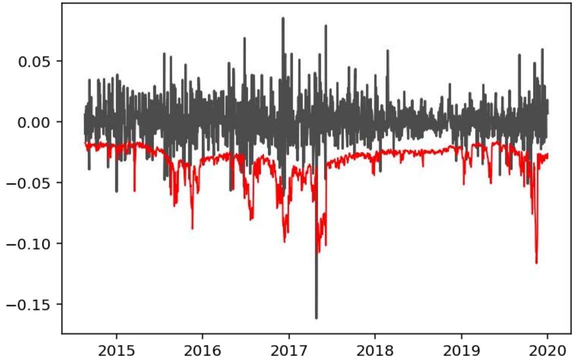
I:ISP - Actual number of exceptions: 94



E:SAN - Actual number of exceptions: 80



H:INGA - Actual number of exceptions: 42



AXYON.AI

Further research

Improving multivariate RCGAN

- Try other methods to determine edges in the graph, e.g. the model could learn the adjacency matrix by itself
- Improve the model scalability, e.g. through weight sharing for the Generator and/or Discriminator
- Estimate the VaR of a portfolio and compare it to multivariate GARCH model

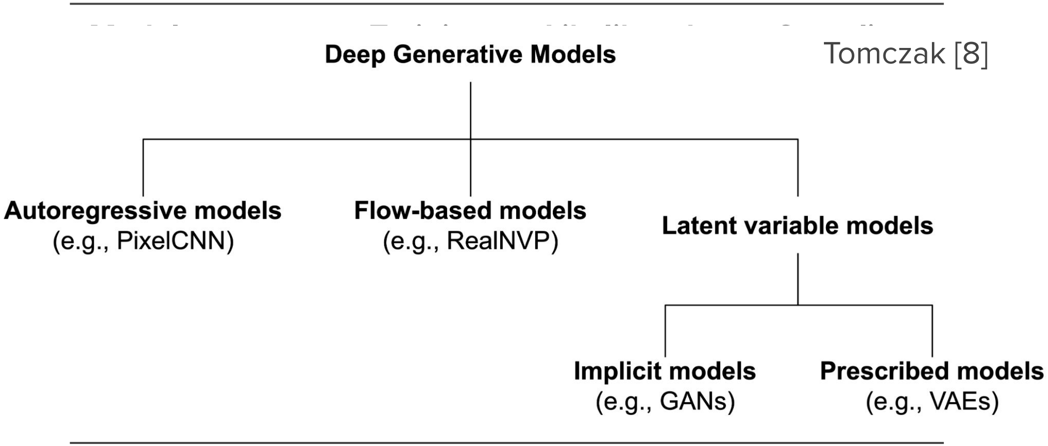
Further research

Beyond GANs

Another future research direction will entail experimenting with different (deep) generative models, evaluating their advantages/disadvantages, and compare them with our SOTA.

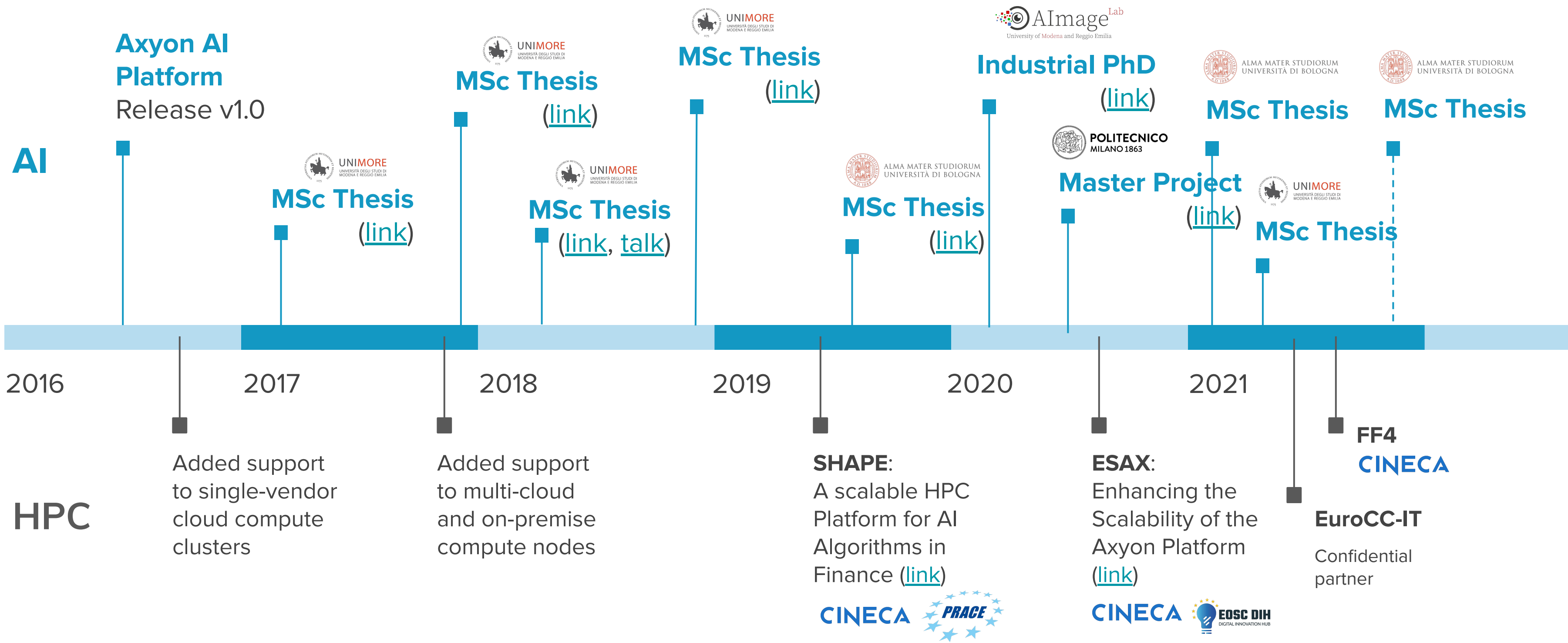
Examples may include:

- Variational Autoencoders (e.g. VRNN, SRNN)
- Flow-based Models (e.g. RealNVP, GLOW)
- Other (Energy-based Models, Diffusion Models, etc.)



Tomczak [8]

Research at Axyon



Acknowledgements

This work was supported by the FF4EuroHPC: HPC Innovation for European SMEs, Project Call 1. The FF4EuroHPC project has received funding from the European High-Performance Computing Joint Undertaking (JU) under grant agreement No 951745.

Resources

[1] Karras, Tero, et al. "Analyzing and improving the image quality of stylegan." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2020.

[2] Goodfellow, I., et al. "Generative adversarial nets." Advances in neural information processing systems. 2014.

[3] Zhang, Daniel, et al. "The ai index 2021 annual report." arXiv preprint arXiv:2103.06312 (2021).

[4] Fiechtner L. ”Risk Management with Generative Adversarial Networks”, MSc Thesis University of Oxford (2019)

[5] Campbell, S. “ A review of backtesting and backtesting procedures”. The Journal of Risk,9(2), 1–17. 2007.

[6] Kipf, T. N., Welling M. ”Semi-supervised classification with graph convolutional networks.”, arXiv:1609.02907 (2016)

[7] Rocheteau E., Tong C. et al. ”Predicting Patient Outcomes with Graph Representation Learning”, arXiv:2101.03940v1 (2021)

[8] Tomczak, Jakub. “Introduction to deep generative modeling: Why, Where and How” https://jmtomczak.github.io/blog/1/1_introduction.html

AXYON.AI

Q&A

Contacts

axyon.ai
via Rainusso 144
Modena, Italy
+39 059 7880067

Jacopo Credi
Co-founder and VP of AI
jacopo.credi@axyon.ai

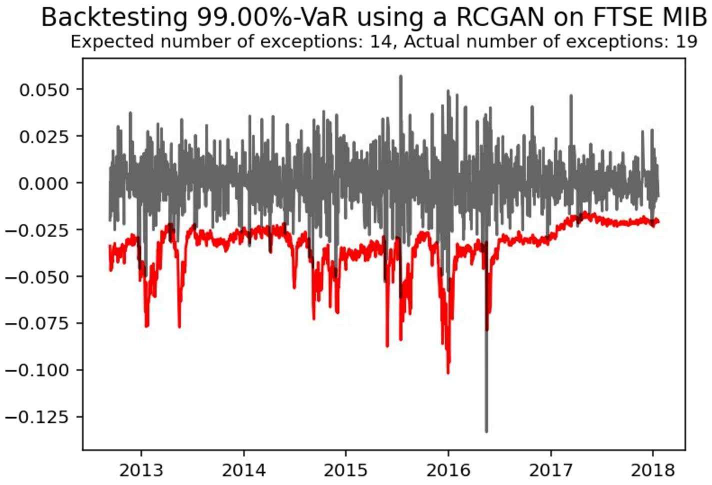
Mirko Savasta
Quantitative Analyst
mirko.savasta@axyon.ai

Giovanni Davoli
Junior Data Scientist
giovanni.davoli@axyon.ai

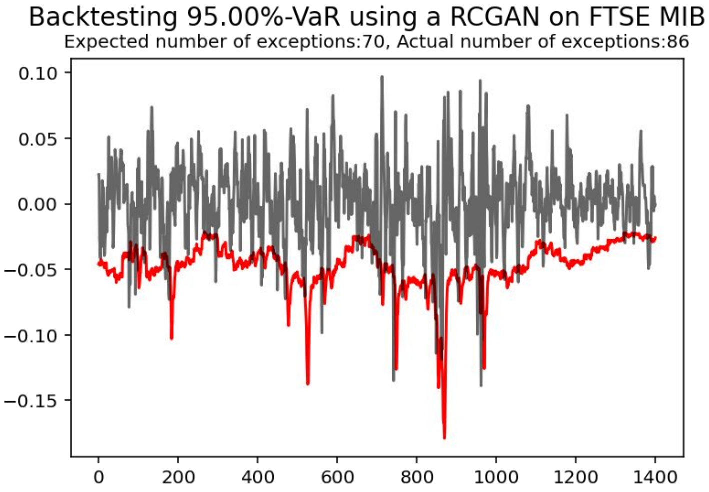
Backup slides

AXYON.AI

Appendix: other VaR estimation with CGAN



We can also estimate the VaR at other confidence levels, e.g. 99% (an important level for bank regulations)



We also tried training the model with 5-day (weekly) returns.

AXYON.AI