Torino

MIDAS – ECML PKDD 2023

September 22, 2023

Comparing Deep RL and Traditional Financial Portfolio Methods

E. Benhamou, JJ Ohana, B. Guez, D. Saltiel, R. Laraki, J. Atif

For Alpha

Dauphine Université Paris | PSL

LAMSADE UMR 7243

Portfolio allocation

- In asset management, there is a gap between mainstream used methods and new machine learning techniques around RL

- DRL has achieved strong results in challenging tasks

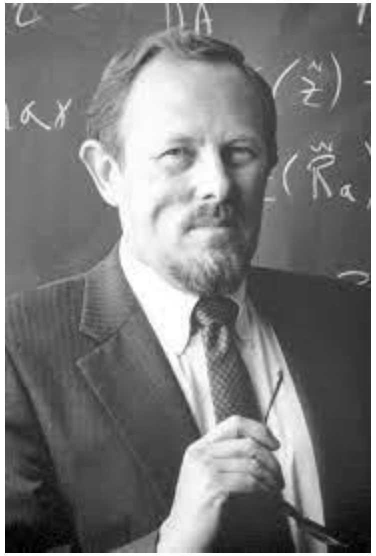


2018 2020 2022 2023

Traditional methods

- Portfolio: Markowitz (1952), and various extensions: Minimum variance, Maximum diversification, Maximum decorrelation, Risk Parity

- In terms of ranking: Sharpe Ratio (1966),



These are indeed optimization

Markowitz, H. 1952.

denote by $w = (w_1, ..., w_l)$ the allocation weights

$\mu = (\mu_1, ..., \mu_l)^T$ be the expected returns

Σ the matrix of variance covariances

r_{min} be the minimum expected return

$$\underset{w}{\text{Minimize}} \quad w^T \Sigma w \qquad (1)$$

$$\text{subject to} \quad \mu^T w \geq r_{min}, \sum_{i=1...l} w_i = 1, 1 \geq w \geq 0$$

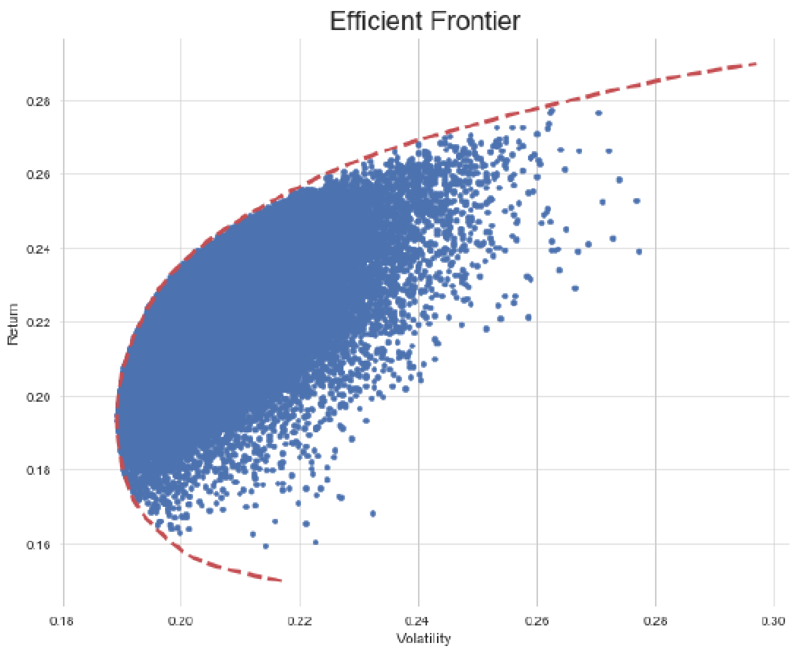


Figure 1: Markowitz efficient frontier for the GAFA: returns taken from 2017 to end of 2019

All methods are indeed convex optimization

Minimum variance portfolio

$$\underset{w}{\text{Minimize}} \quad w^T \Sigma w$$
$$\text{subject to} \quad \sum_{i=1\ldots l} w_i = 1, 1 \geq w \geq 0$$

Chopra and Ziemba 1993

Maximum diversification portfolio

$$\underset{w}{\text{Maximize}} \quad \frac{w^T \sigma}{\sqrt{w^T \sum w}}$$
$$\text{subject to} \quad \sum_{i=1\ldots l} w_i = 1, 1 \geq w \geq 0$$

$\sigma = (\Sigma_{i,i})_{i=1..l}$ the diagonal elements of the covariance matrix Σ

Choueifaty and Coignard 2008

What about RL?

$$\underset{\pi(.)}{\text{Maximize}} \quad \mathbb{E}[R_T]$$

A more general objective

$$\text{subject to} \quad a_t = \pi(s_t)$$

With general constraints

where R R_T cumulative reward to be defined later and

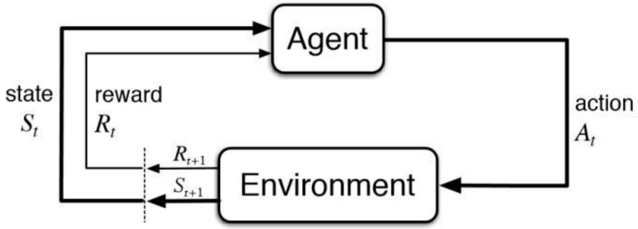
$a_t = \pi(s_t)$ he action given by a policy that is a function of states

(more to come)

Why RL and not supervised?

Reinforcement Learning (RL):

RL consists in **finding the optimal action** A_t^* (the portfolio allocation) **according to states** S_t (financial information) **given a reward** R_t (the best net portfolio final performance)



$$A_t^* = \pi(S_t)$$

RL **learns** and **finds the optimal portfolio allocation** in an interactive environment **by trial and error** using feedback from actions and rewards

Supervised Learning (SL):

Quite general and encompasses **classification and regression.** Goal is to **infer a function** from labelled training data that maps inputs into outputs

we observe $D_n = \{(X_1, Y_1), \ldots, (X_n, Y_n)\}$

n independent random copies of $(X, Y) \in \mathcal{X} \times \mathcal{Y}$

find a function $f : \mathcal{X} \mapsto \mathbf{R}$

f that minimizes the expected ℓ-risk

$$\mathcal{R}_\ell(h) = \mathbf{E}_{X \times Y}[\ell(Y, f(X))].$$

Observations

- Regular observations:
  - Past returns $r_t = \frac{p_t^k}{p_{t-1}^k} - 1$ where p_t^k is the price at time t of the asset
  - Empirical standard deviations useful to detect regime changes $\sigma_t^k = \sqrt{\frac{1}{d}\sum_{u=t-d+1}^{t}(r_u - \mu)^2}$

→ three dimensional tensor $A_t = [A_t^1, A_t^2]$

with $A_t^1 = \begin{pmatrix} r_{t-i_j}^1 & \dots & r_t^1 \\ \dots & \dots & \dots \\ r_{t-i_j}^m & \dots & r_t^m \end{pmatrix}$, $A_t^2 = \begin{pmatrix} \sigma_{t-i_j}^1 & \dots & \sigma_t^1 \\ \dots & \dots & \dots \\ \sigma_{t-i_j}^m & \dots & \sigma_t^m \end{pmatrix}$

- Contextual information:
  - Equity data
  - Fixed income data
  - Credit data
  - Interactions between variables

→ Two-dimensional tensor

$$C^t = \begin{pmatrix} c_t^1 & \dots & c_{t-i_k}^1 \\ \dots & \dots & \dots \\ c_t^p & \dots & c_{t-i_k}^p \end{pmatrix}$$

What about the different RL Algorithms?

Comparison of DRL algorithms							
Algorithms	Input	Output	Type	State-action spaces support	Finance use cases support	Features and Improvements	Advantages
DQN	States	Q-value	Value based	Discrete only	Single stock trading	Target network, experience replay	Simple and easy to use
Double DQN	States	Q-value	Value based	Discrete only	Single stock trading	Use two identical neural network models to learn	Reduce overestimations
Dueling DQN	States	Q-value	Value based	Discrete only	Single stock trading	Add a specialized dueling Q head	Better differentiate actions, improves the learning
DDPG	State action pair	Q-value	Actor-critic based	Continuous only	Multiple stock trading, portfolio allocation	Being deep Q-learning for continuous action spaces	Better at handling high-dimensional continuous action spaces
A2C	State action pair	Q-value	Actor-critic based	Discrete and continuous	All use cases	Advantage function, parallel gradients updating	Stable, cost-effective, faster and works better with large batch sizes
PPO	State action pair	Q-value	Actor-critic based	Discrete and continuous	All use cases	Clipped surrogate objective function	Improve stability, less variance, simply to implement
SAC	State action pair	Q-value	Actor-critic based	Continuous only	Multiple stock trading, portfolio allocation	Entropy regularization, exploration-exploitation trade-off	Improve stability
TD3	State action pair	Q-value	Actor-critic based	Continuous only	Multiple stock trading, portfolio allocation	Clipped double Q-Learning, delayed policy update, target policy smoothing.	Improve DDPG performance
MADDPG	State action pair	Q-value	Actor-critic based	Continuous only	Multiple stock trading, portfolio allocation	Handle multi-agent RL problem	Improve stability and performance

Objective

- Action: long only allocation that sum to 100% in 11 assets :
  - **Equity indexes** (4) : S&P 500, Eurostoxx 50, Nikkei 225, FTSE 100.
  - **Bonds** (4): US 10-year TNote, European Bund, UK 10-year Gilt, Japanese Government Bond 10-year.
  - **Commodities** (3): Brent Oil, Gold, and Copper

- Reward: Get the highest Sharpe out of sample

Features

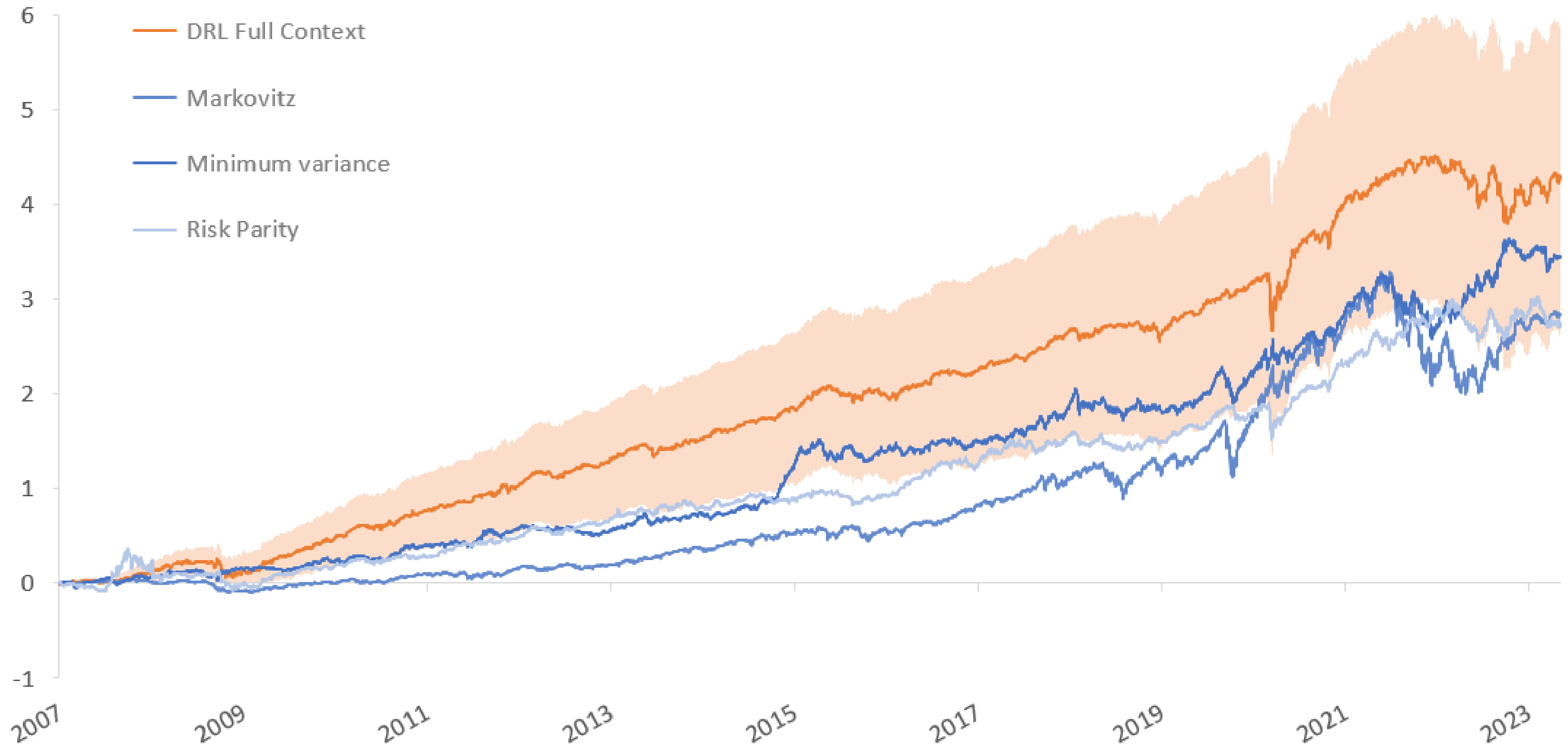
Table 1. Regular observations. List of all portfolio assets

Bloomberg Ticker	Description	Category
ES1 Index	E-Mini S&P 500 Futures	Equity
NH1 Index	Nikkei 225 Futures (Yen)	Equity
VG1 Index	Euro Stoxx 50 Futures	Equity
Z 1 Index	FTSE 100 Futures	Equity
TY1 Index	10Y T-Note Futures	Rates
RX1 Index	Euro-Bund Futures	Rates
G1 Index	ICE Long Gilt Futures	Rates
JB1 Index	10Y JGB Futures	Rates
GC1 Index	Gold Futures	Commodity
CO1 Index	Brent Crude Futures	Commodity
HG1 Index	High Grade Copper Futures	Commodity

Table 2. List of all the features used as contextual variables

Bloomberg Ticker	Description	Category
USGG10YR Index	US Rates 10 year	Interest Rates
USGG2YR Index	US Rates 2 year	Interest Rates
GDBR10 Index	Europe Rates 10 year	Interest Rates
GDBR2 Index	Europe Rates 2 year	Interest Rates
VIX INDEX	Cboe Volatility Index	Market Sentiment
CDX HY CDSI GEN 5Y SPRD Corp	MARKIT Credit CDX HY index	Credit Default Swaps
ITRX XOVER CDSI GEN 5Y Corp	MARKIT Credit Europe Itraxx Crossover	Credit Default Swaps
DXY Index	Dollar index	Currency Rates
PFRGGDP6 Index	Survey of Professional Forecaster on GDP	Economic Forecast
PIIFFCPI1 Index	Survey of Professional Forecaster on CPIT	Economic Forecast
DOESTCRD Index	Crude Oil Total Inventory	Commodity Market
COMXGOLD Index	Comex Gold Inventory Data	Commodity Market
COMXCOPR Index	Comex Copper Inventory Data	Commodity Market
CESIGL Index	Global Economic Surprise	Economic Indicators
CESIUSD Index	USD Economic Surprise	Economic Indicators
CESIEUR Index	EUR Economic Surprise	Economic Indicators
CESIJPY Index	JPY Economic Surprise	Economic Indicators
CESIEM Index	EM Economic Surprise	Economic Indicators

Results



	DRL Mean	DRL Max	DRL Min	Markovitz	Minimum variance	Risk Parity
Ann. return	10.7%	12.6%	8.3%	8.6%	9.6%	8.4%
Ann. Volatility	7.1%	5.8%	9.6%	12.7%	8.8%	10.0%
Sharpe ratio	1.50	2.17	0.86	0.67	1.09	0.83
Max DD	16.4%	13.3%	21.7%	30.0%	16.0%	31.9%
Calmar ratio	0.65	0.95	0.38	0.29	0.60	0.26

Conclusion

- Deep learning and reinforcement techniques show great potential in financial portfolio allocation as it:
  - Can extract valuable insights from intricate financial data for better investment decisions.
  - Can adapt behavior to context

- Traditional financial techniques can be formulated as equivalent reinforcement learning problems.

- An experiment confirmed DRL outperformed conventional methods :
  - Higher annual returns
  - Superior risk-adjusted returns (Sharpe ratio)
  - Better management of maximum drawdowns.