Deep Reinforcement learning for stock portfolio allocation

Ismail Benaija
Taha Boukhari Daniel Lopez Herranz

14 April 2022

Table of Contents

Context

Traditional approach:

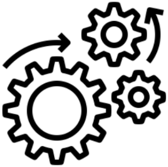
Trade off between Return and Risk

One-step optimization problems:

- Min Variance
- Maximum Decorrelation portfolio
- Maximum Diversification portfolio
- Risk parity portfolio

Context

Some Machine Learning use cases in Finance:

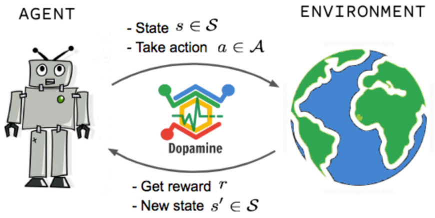


Process automation

Fraud detection

Algorithmic trading

Problematic

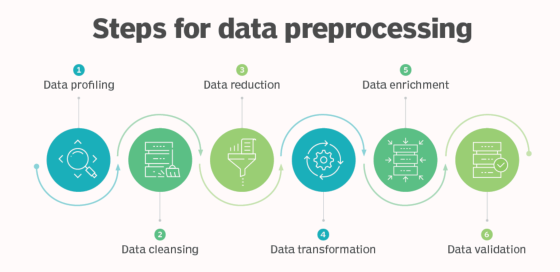


- Automated trading solution for portfolio allocation
- Build a DRL agent and explain its decisions

Data

- *yfinance* python module
- Stock : 30 stock of DOW JONES
- Start date : *2008-01-01*
- End date : *2021-09-02*

Data preprocessing



- Build covariance matrix
- Compute technical indicators : macd rsi, atr, dx
  - Time frame 5 days
  - Time frame 30 days

Technical indicators

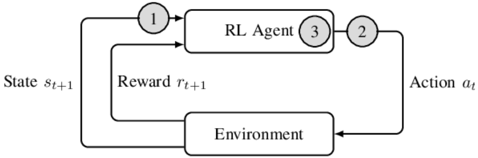
RSI : It is a momentum indicator that measures the speed and change in the movement of a price.

MACD : It is a trend indicator that is equal to the difference between the 12 day exponential moving average and the 26 day one.

DX : is an intermediate result to evaluate the strength of a trend and to define a period of sideway trading.

ATR : It is a volatility indicator that indicate the difference between today's high and today's low in an asset price.

RL Introduction



- The agent acts in an environment, observing its state and receiving rewards
- From its perceptual and reward information, it must determine what to do

Gym Framework

OpenAI

- Toolkit for developing and comparing reinforcement learning algorithms
- Supports building a custom environment. For that it's necessary to define:
  - Observation and Action Spaces
  - Rewards
  - *step* function
  - *reset* function

Custom Environment

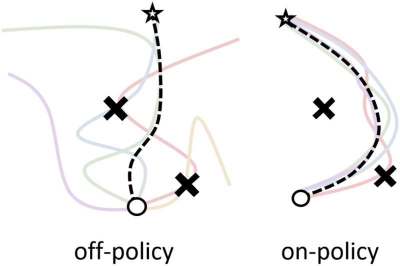
- Observation Space:
  - Covariance Matrix
  - Technical Indicators
- Action Space:
  - Allocation for each stock
  - $Box(0, 1, N)$
- Reward:
  - Portfolio value variation
  - Total portfolio value

Environment Parameters

Main parameters

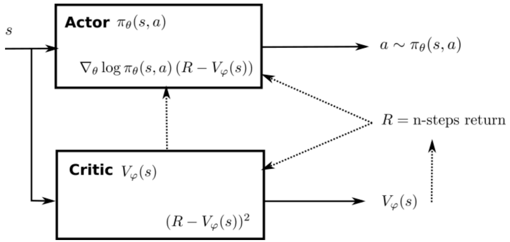
- *tech indicator list*
- *transaction cost pct*
- *hmax*
- *reward scaling*
- ...

Proximal Policy Optimization (PPO)



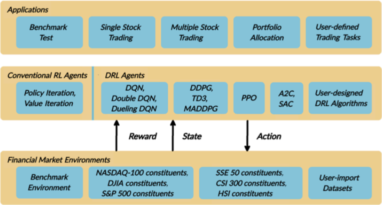
- Used to control updates to the policy gradient to ensure that the new policy is better
- Simple to use and provide very good results overall

Synchronous Advantage Actor Critic (A2C)



- Combines two types of models
- Learning which states are better or worse
- Teach the agent to seek out good states and avoid bad states

FinRL



- Three principal layers :
  - Environment
  - Agent
  - Application

Backtesting

Period of the backtest: July 2020-September 2021

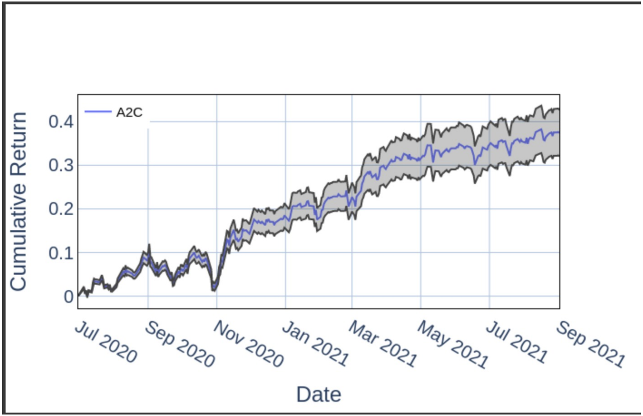
Algorithm	Initial portfolio	Final portfolio	Sharpe
A2C	1 000 000	1 448 550	2.19
PPO	1 000 000	1 446 500	2.21

- Very similar results
  - A2C returns are better
  - PPO Sharpe ratio is better

Algorithm comparison

- Cummulative annual return :
  - A2C : 44%
  - PPO : 44%
  - SVM : 25%

Confidence Interval



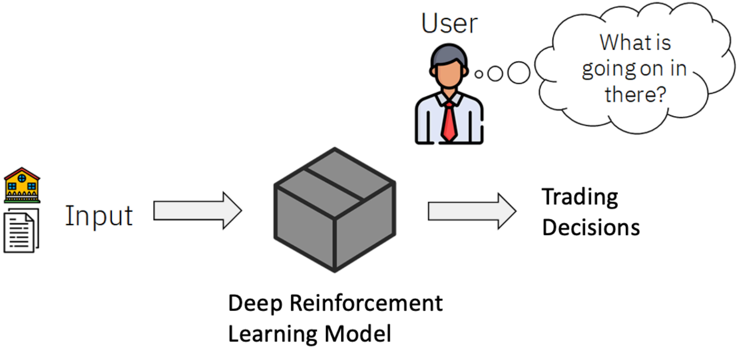
Actions for each sector

We can see the distribution of action of every sector within 2 days

date	Healthcare	Technology	Industrials	Energy	Financial Services	Consumer	Communication Services
2020-07-01	0.185185	0.222222	0.148148	0.037037	0.111111	0.222222	0.074074
2020-07-02	0.159510	0.291676	0.125383	0.064539	0.072369	0.224049	0.062474

Explainability

Explainable AI is a field that aims to make the decisions of an model understandable by the human.



Explainability

It is an important step when using algorithms for trading in finance because of:

Accountability:
High stakes with high amounts of money in play

Fairness:
Important decisions (eg: declining credit card)

Transparency:
On what information is the decision based

Explainability

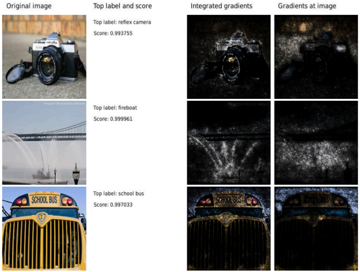
There are two big type of explainability techniques:

- Post-Hoc: Build an explainability layer on top of the model to extract insights like feature importance (eg: Shapley)
- Transparent model: Use the model directly (eg: Decision Trees)

We choose to focus on feature importance especially technical indicators because it's a set of features that capture the trend, the momentum, volatiliy in a financial asset and that are used for trading strategies.

Explainability

Our goal is to measure which technical indicator contributed the most to a reallocation in the portfolio. For that, we will use a perturbation technique: Integrated Gradients. It is an explainability technique for deep neural networks which visualizes its input feature importance that contributes to the model's prediction through a saliency map.



Explainability

The formula for the IG is the following:

$$IntegratedGradients_i(x) ::= (x_i - x'_i) \times \int_{\alpha=0}^{1} \frac{\partial F(x' + \alpha \times (x - x'))}{\partial x_i} d\alpha$$

where:

$_i$ = feature
x = input
x' = baseline
α = interpolation constant to perturb features by

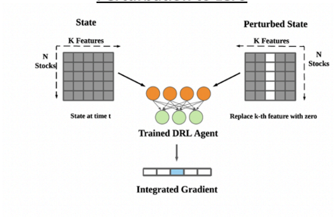
Therefore the type of perturbation that is choosen is important

1. Context and problematic
2. Trading Agent
3. **Explainability**
   - Introduction
   - Method: Integrated Gradients
   - **Implementation**
   - Results
4. Discussion and limits
5. References

Explainability

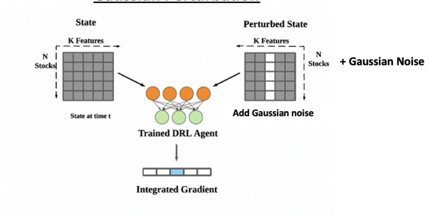
We decided to experiment two types of perturbations that suited best the stock portfolio allocation situation:

Perturbation to zero



Interesting in the context of technical indicators as features because they don't have the same space of values

Gaussian Perturbation



Classic type of perturbation

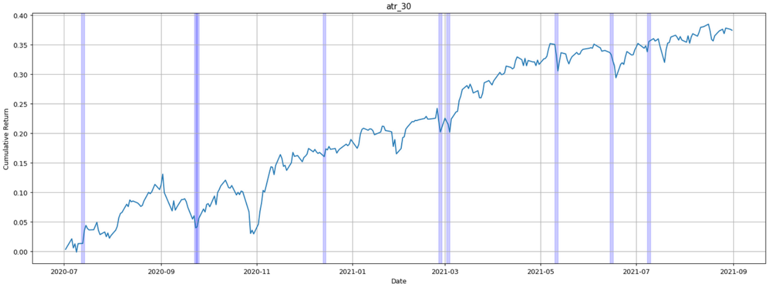
Explainability

An example of explainability results with PPO and a random perturbation

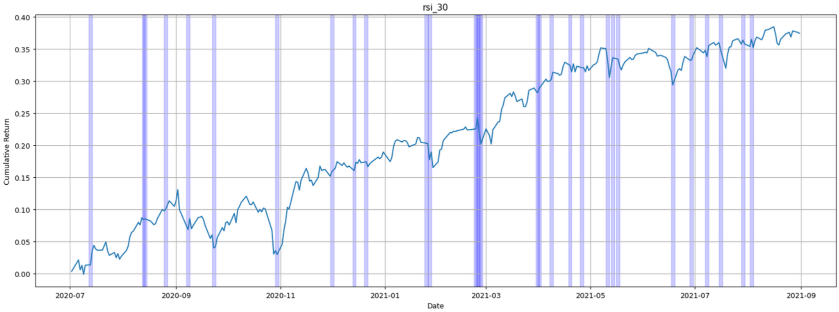
	macd	rsi 30	cci 30	dx 30
min	-0.00013	-0.0028	-0.0024	-0.00085
max	0.00038	92.2	0.0015	0.0072
mean	0.000001	5.19	-0.000027	0.00033
std	0.00001	13.2	0.00024	0.0093

Explainability

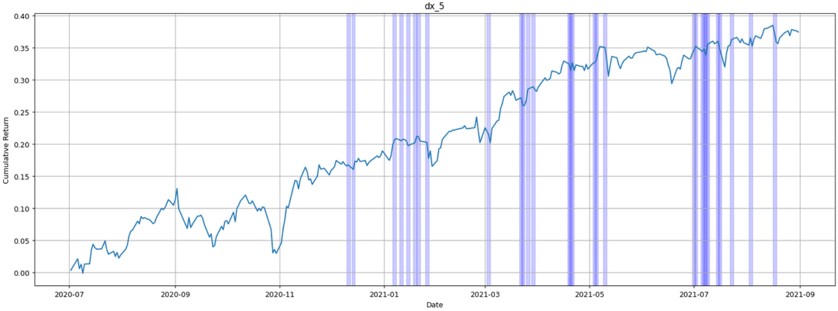
A way to better visualize these values is to look at the 0.9 quantile for each indicator and plot "detection zones"



Explainability



Explainability



Explainability

Finally, a way for a potential user to benefit from this explainability module will be some natural language sentences on each step of trading highlighting key points:

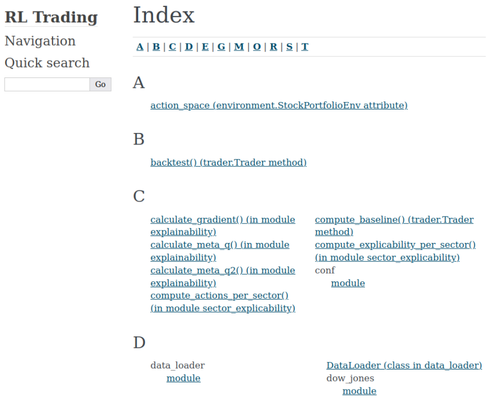
- Total Reallocation.
- Technical indicator that impacted this change the most and its variation.
- The stock that had the biggest reallocation with the technical indicator that caused it and its variation.

```
Quel jour tu veux regarder?: (au format AAAA-MM-JJ)
2020-07-06
L'allocation a changé de 35%
L'indicateur technique qui a le plus influencé ce changement est rsi_5. Celui-ci a varié en moyenne pour tout les s
tocks de 17% entre hier et aujourd'hui
Le stock qui a la plus grande réallocation est : MCD
Cette réallocation est influencé par l'indicateur technique suivant : rsi_30. Celui-ci a varié pour ce stock de 6%
```

Discussion and limits

- We have only considered technical indicators individually but it would probably be more interesting to look for a combination of indicators and see how they react together.
- It could be the case that technical indicators play no role on the agent decision and that our agent only use the price action of the stocks.

Discussion and limits



References

1. - Title : HEDGING OF FINANCIAL DERIVATIVE CONTRACTS VIA MONTE CARLO TREE SEARCH
   - Author : OLEG SZEHR
2. - Title: Explain Your Move: Understanding Agent Actions Using Specific and Relevant Feature Attribution
   - Author: Nikaash Puri, Piyush Gupta et al.
   - Date: 2019
3. - Title: Asynchronous Methods for Deep Reinforcement Learning
   - Author: Volodymyr Mnih, Adrià Puigdomènech Badia
   - Date: 2016
4. - Title: Proximal Policy Optimization Algorithms
   - Author: John Schulman, Filip Wolski, Prafulla Dhariwal
   - Date: 2017