Portfolio Management using Deep Reinforcement Learning

Huan-Rui Zhang

Background

- **Goal**: Build up a DDPG agent to maximize the profit in Vanguard S&P 500 Exchange Traded Fund (ETF) (VOO)

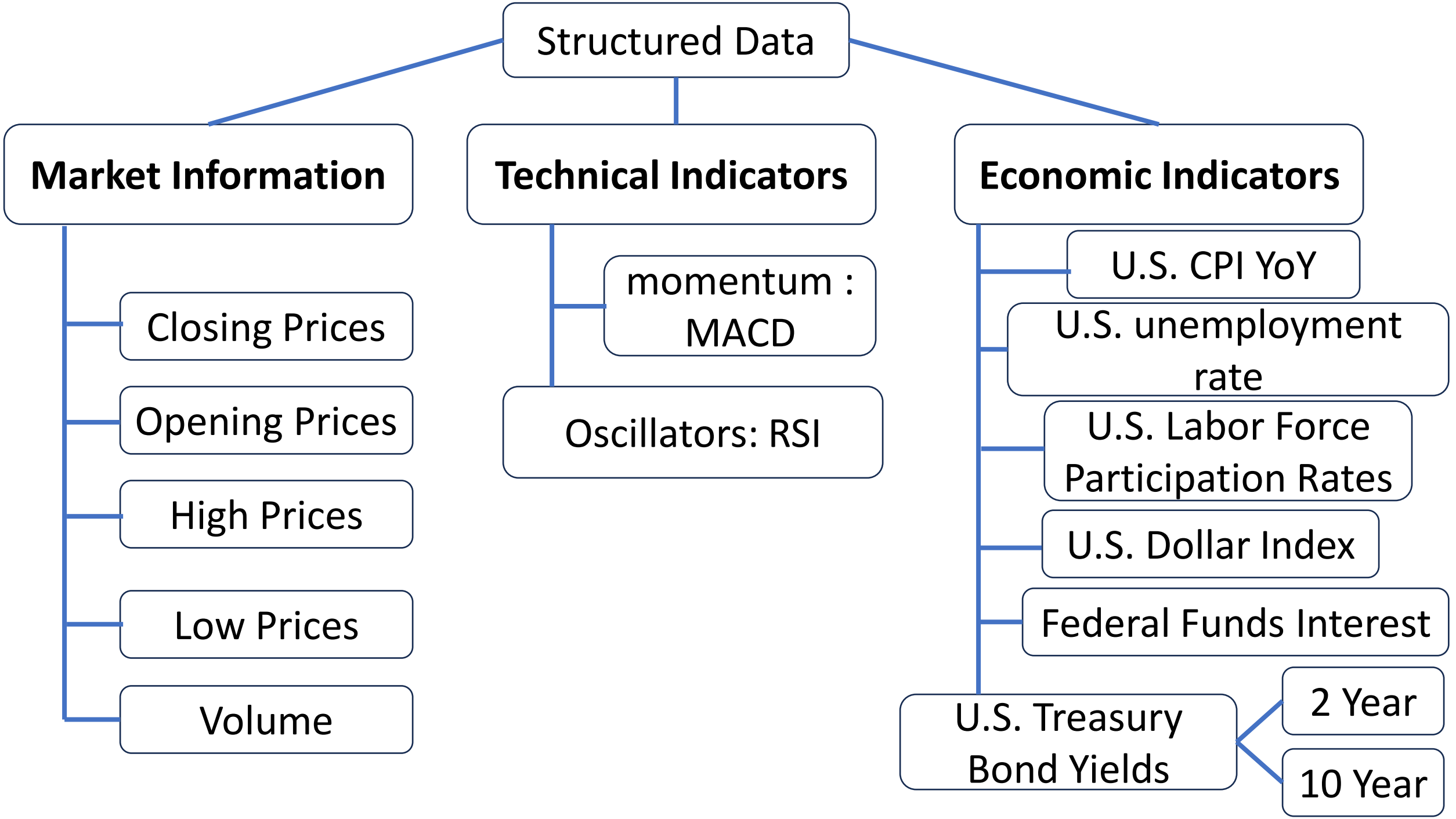
- **Assumptions**:
  - The liquidity of VOO is high enough that the agent can carried at the last price when an action is taken.
  - The investment made by this agent has no influence on the market.
  - Zero trading fee
  - Zero expense ratio (0.03% in real world)

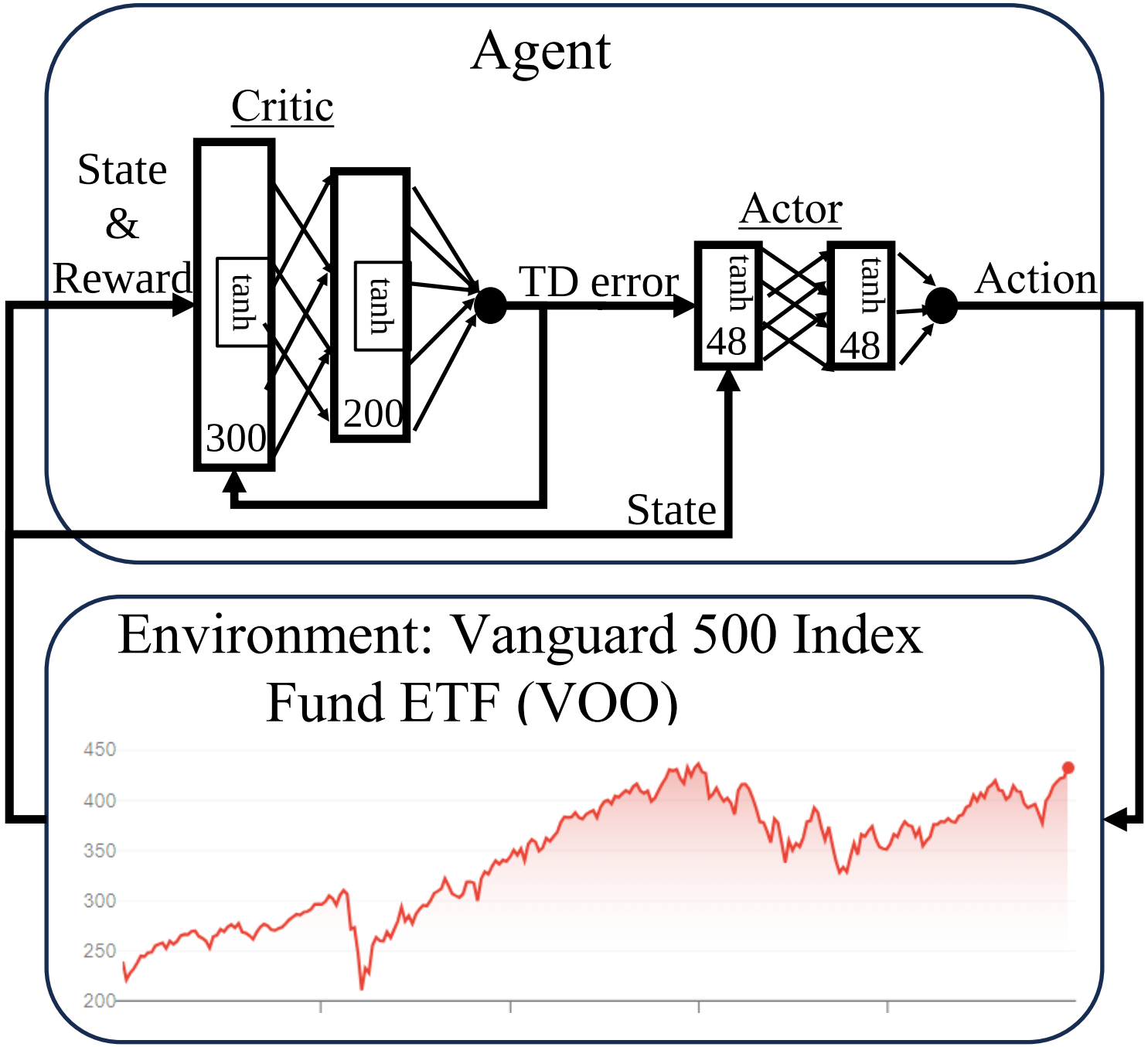
- Structured Data related to stock market (Next slide)

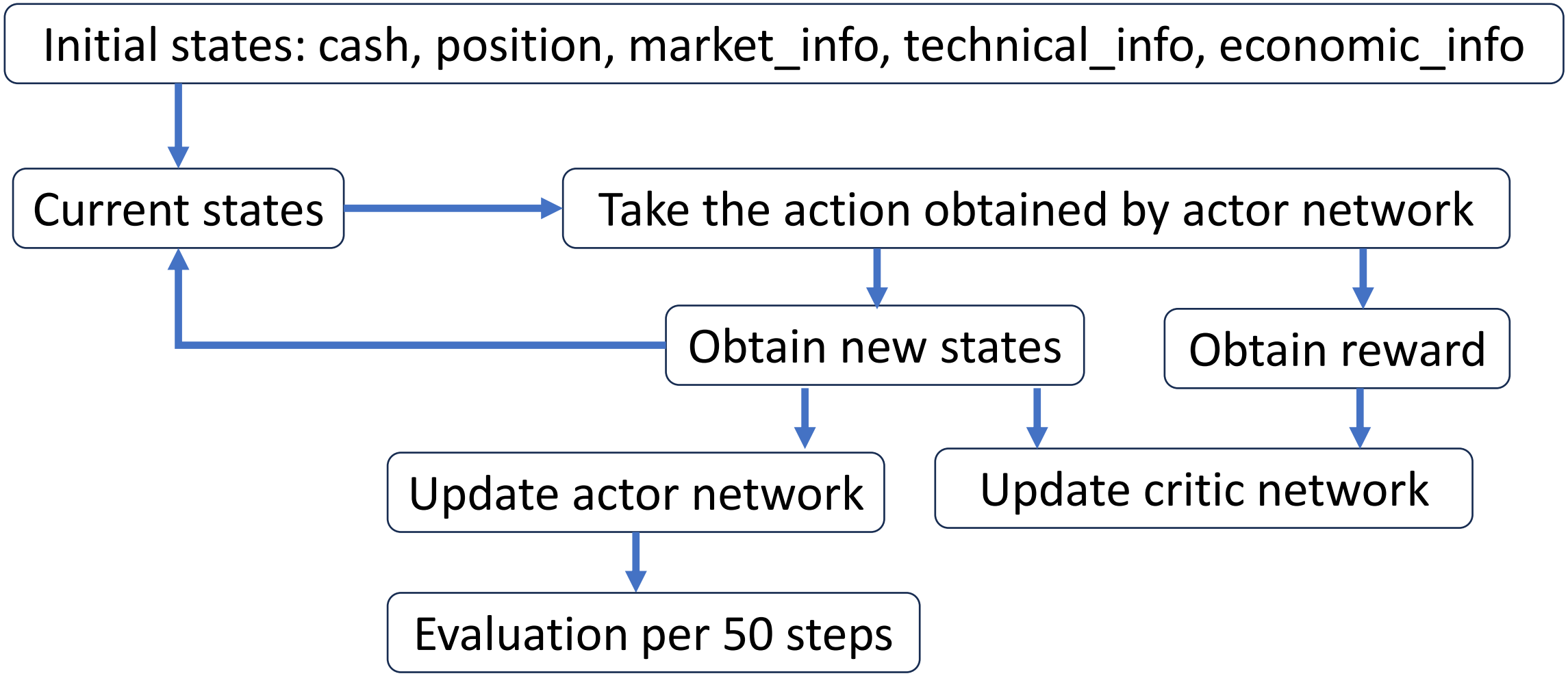
MDP Formulation (1/2)

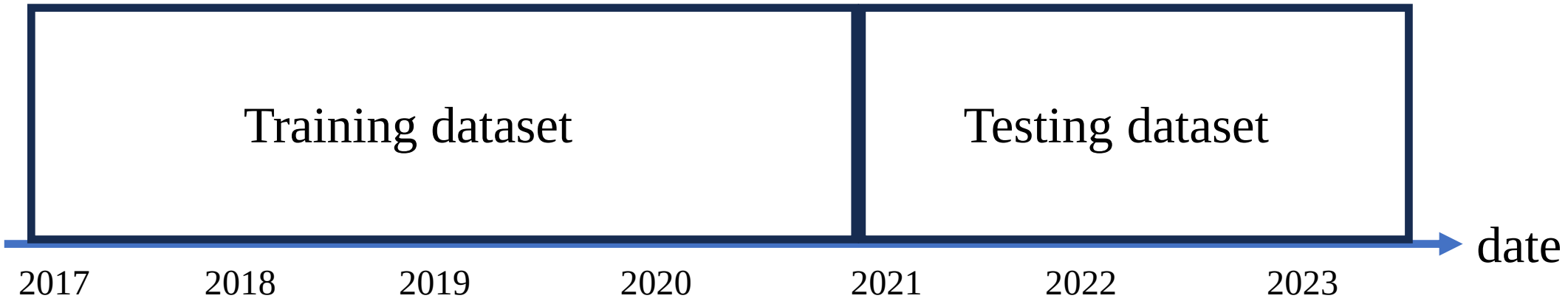
- **Agent**: A trading software
- **Environment**: Vanguard 500 Index Fund ETF (VOO).
- **State** (All in continuous space):
  - Cash: 0 ~ 10000 USD
  - Position: 0 ~ 20

  - **Market information**: closing price: 200 ~ 450
  - **Technical indicators**: MACD, RSI
  - **Economic indicators**: CPI YoY, US unemployment rate, U.S. Dollar Index, Fed Fund Interest, US 2-year and 10-year Treasury Bond Yields


Structured Data
Market Information
Technical Indicators
Economic Indicators
Closing Prices
Opening Prices
High Prices
Low Prices
Volume
momentum : MACD
Oscillators: RSI
U.S. CPI YoY
U.S. unemployment rate
U.S. Labor Force Participation Rates
U.S. Dollar Index
Federal Funds Interest
U.S. Treasury Bond Yields
2 Year
10 Year


Agent
Critic
State
&
Reward
tanh
300
tanh
200
TD error
Actor
tanh
48
tanh
48
Action
State
Environment: Vanguard 500 Index
Fund ETF (VOO)
450
400
350
300
250
200


Initial states: cash, position, market_info, technical_info, economic_info
Current states
Take the action obtained by actor network
Obtain new states
Obtain reward
Update actor network
Update critic network
Evaluation per 50 steps


Training dataset
Testing dataset
date
2017
2018
2019
2020
2021
2022
2023

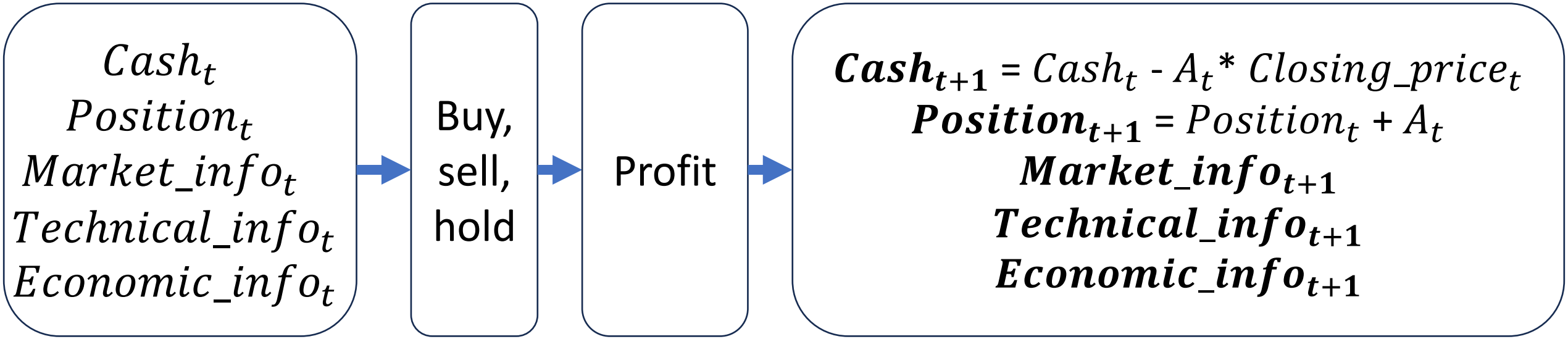
MDP Formulation (2/2)

- **Actions**: sell, hold, buy in continuous space -20 ~ 20
- **Reward**:

Profit = $Cash_{t+1}$ + ($Position_{t+1}$ * $Closing\_price_t$)

– Original_cash - (Original_position * Original_price)

- **Transition dynamics**: $P_r(S_{t+1}, R_{t+1} \mid S_t, A_t)$



Problem Classification

- Continuous task
  - There is no terminal state for stock market in this application.

- Deterministic policy
  - Learn deterministic policy from experience without knowing the system dynamics and policy.

Suggested Solution Algorithm: DDPG

- Strengths
  - Continuous and very large action spaces
  - Addresses maximization bias
  - TD methods suitable for continuous task

- Drawback
  - local optimum

References

[1] O Bustos, A. Pomares-Quimbaya, "Stock market movement forecast: A Systematic review," *Expert Systems with Applications*, Volume 156,

2020

[2] Htun HH, Biehl M, Petkov N. Survey of feature selection and extraction techniques for stock market prediction," *Financ Innov. 2023*;9(1):26.

[3] Nti, I.K., Adekoya, A.F. & Weyori, B.A. "A systematic review of fundamental and technical analysis of stock market predictions," Artif Intell Rev 53, 3007–3057 (2020)

[4] Alhomadi, Abraham, "Forecasting Stock Market Prices: A Machine Learning Approach" *(2021). All Graduate Plan B and other Reports*. 1610

[5] Y. -J. Hu and S. -J. Lin, "Deep Reinforcement Learning for Optimizing Finance Portfolio Management," *2019 Amity International Conference on Artificial Intelligence (AICAI)*, Dubai, United Arab Emirates, 2019, pp. 14-20

[6] Mohammad Amin Masoudi, "Robust Deep Reinforcement Learning for Portfolio Management," *Master's degree in Master of Science in Management in Telfer School of Management, University of Ottawa, 2021*