DEEP REINFORCEMENT LEARNING FOR OPTIMAL PORTFOLIO MANAGEMENT

The Literature

Outline

- *DQN*
  - *Deep Reinforcement Learning for Trading (Zhang et al., 2019)*
  - *A Deep Q-learning Portfolio Management Framework for the Cryptocurrency Market (Lucarelli et al., 2019)*
  - *Application of Deep Q-Network in Portfolio Management (Gao et al., 2020)*
- *DDPG*
  - *Adversarial Deep Reinforcement Learning in Portfolio Management (Liang et al., 2018)*
  - *Practical Deep Reinforcement Learning Approach for Stock Trading (Xiong et al., 2018)*
  - *Deep Reinforcement Learning for Portfolio Management based on the Empirical Study of Chinese Stock Market (Huang et al., 2021)*

Deep Reinforcement Learning for Trading (Zhang et al., 2019)

- Dueling Double DQN (based on LSTM architecture)
- Discrete action space ⟶ discretization of weight space for assets
- 50 ratio-adjusted futures contracts for commodity, equity index, fixed income and FX asset classes
  - Different models trained for each asset class
- *Volatility Scaling* to scale trade positions based on market volatility
- Evaluated on 9 different metrics
- DQN outperforms baseline strategies (Long, Momentum, MACD) and 2 other RL algorithms (PG, A2C) for all asset classes except equity

A Deep Q-learning Portfolio Management Framework for the Cryptocurrency Market (Lucarelli et al., 2019)

- DQN, D-DQN, DD-DQN
- 4 asset portfolio: Bitcoin (BTC), Litecoin (LTC), Ethereum (ETH), Riple (XRP).
- *Action Space: {Buy, Hold, Sell} 30 intervals b/w* [$0.83, $25]
- ε-greedy strategy
- Each portfolio asset represented by *local agent* which competes with others
  - *N deep RL agents instead of 1*
  - *Local Reward Functions: a)* Sum of nominal net returns, and b) Sharpe Ratio
- *Global agent* manages global reward functions
  - Combination of local rewards
- Varying results over different test periods. However, positive average daily returns for all 6 models, and better performance than market baseline

Application of Deep Q-Network in Portfolio Management (Gao et al., 2020)

- DD-DQN (CNN based architecture with dueling Q Net)
- 5 US stock portfolio (CAH, CAT, CCE, CCL, DIS)
- *(Prioritized) Experience Replay* - Samples with greatest error chosen more often
- *Assumptions*: a) No market impact, b) Zero transaction fee, c) Zero Slippage
- Evaluated on ARR and Sharpe Ratio
  - *Risk resistance evaluated using MDD → lowest risk policy obtained*
- Outperforms 10 traditional strategies with 25% higher returns than next best

Adversarial Deep Reinforcement Learning in Portfolio Management (Liang et al., 2018)

- DDPG - Actor (Policy Gradient) + Critic (Q Learning)
  - *4 Networks: Online actor, online critic, target actor, target critic*
  - *CNN with deep residual network to prevent gradient vanishing*
  - *Experience Replay*
- 5 assets randomly picked from asset pool
- *Adversarial Training:* Add random noise to market prices for greater training robustness and risk sensitivity
- Objective function: Risk-adjusted portfolio value
- *Assumptions*: a) Continuous Market, b) Weight reallocations done at close price
- Evaluated on a) Daily average return, and b) Sharpe Ratio
  - Adversarial training compared with baseline
- Outperform *University of California Retirement Plan* (UCRP)

Practical Deep Reinforcement Learning Approach for Stock Trading (Xiong et al., 2018)

- DDPG
- 30 stock portfolio from the Dow Jones
- Evaluated on a) Annualized returns, and b) Sharpe Ratio
- ~50% higher returns compared to index

Deep Reinforcement Learning for Portfolio Management based on the Empirical Study of Chinese Stock Market (Huang et al., 2021)

- DDPG (Modified CNN based architecture)
- Randomly selected portfolios from CSI500
- Uses *{P/E, P/B, Turnover}* factors in addition to *{op, cl, hi, lo}*
- Introduce *shorting* and *arbitraging* in continuous action space
- Objective function: Daily log returns
- *Assumptions:* a) No market impact, b) No slippage
- Evaluated against market (CSI300) on 7 metrics

To-do

- Decisions on:
  - *Network Type(s) and variants*
  - *Hyperparameter Tuning*
  - *Objective Function*
  - *Portfolio Selection*
  - *Discrete Action Space*
  - *Transaction costs?*
- Familiarization with the tools
- Implementations of DQN and DDPG architectures
- More Literature

FIN