Reinforcement Learning-Based Strategies for High-Frequency Trading

Group Number: B455 Group Member: 陳冠菻 Professor: 馬席彬

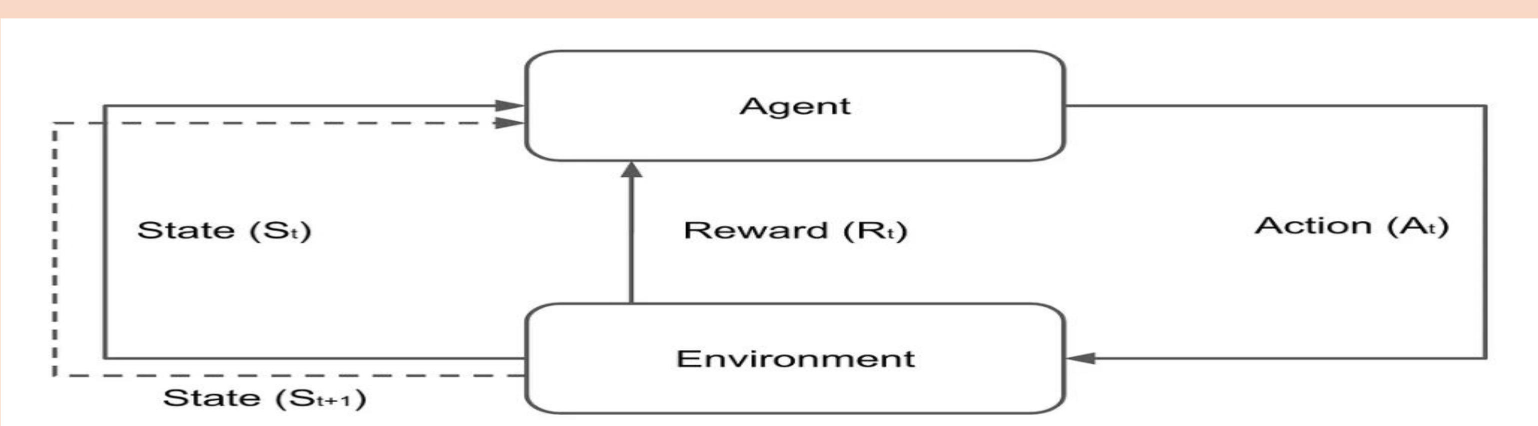
Abstract

With advancements in technology and AI, fintech has flourished. A Q-learning-based high-frequency trading system stands out by learning market fluctuations and optimizing trades in dynamic environments. Using Q-learning, it adapts to market changes, aiming to maximize returns and improve performance while accounting for trading costs.

Overview of Q-Learning

Q-learning is a model-free reinforcement learning algorithm that learns optimal actions in dynamic environments. It updates a Q-value table using the formula:

$$Q(s, a) \leftarrow Q(s, a) + \alpha[r + \gamma \max Q(s', a') - Q(s, a)]$$



System Design

The success of the Q-learning trading system heavily depends on the thoughtful design of its state, action, and reward components. These elements allow the model to effectively interpret market conditions, make decisions, and learn from outcomes.

State, Action, and Reward

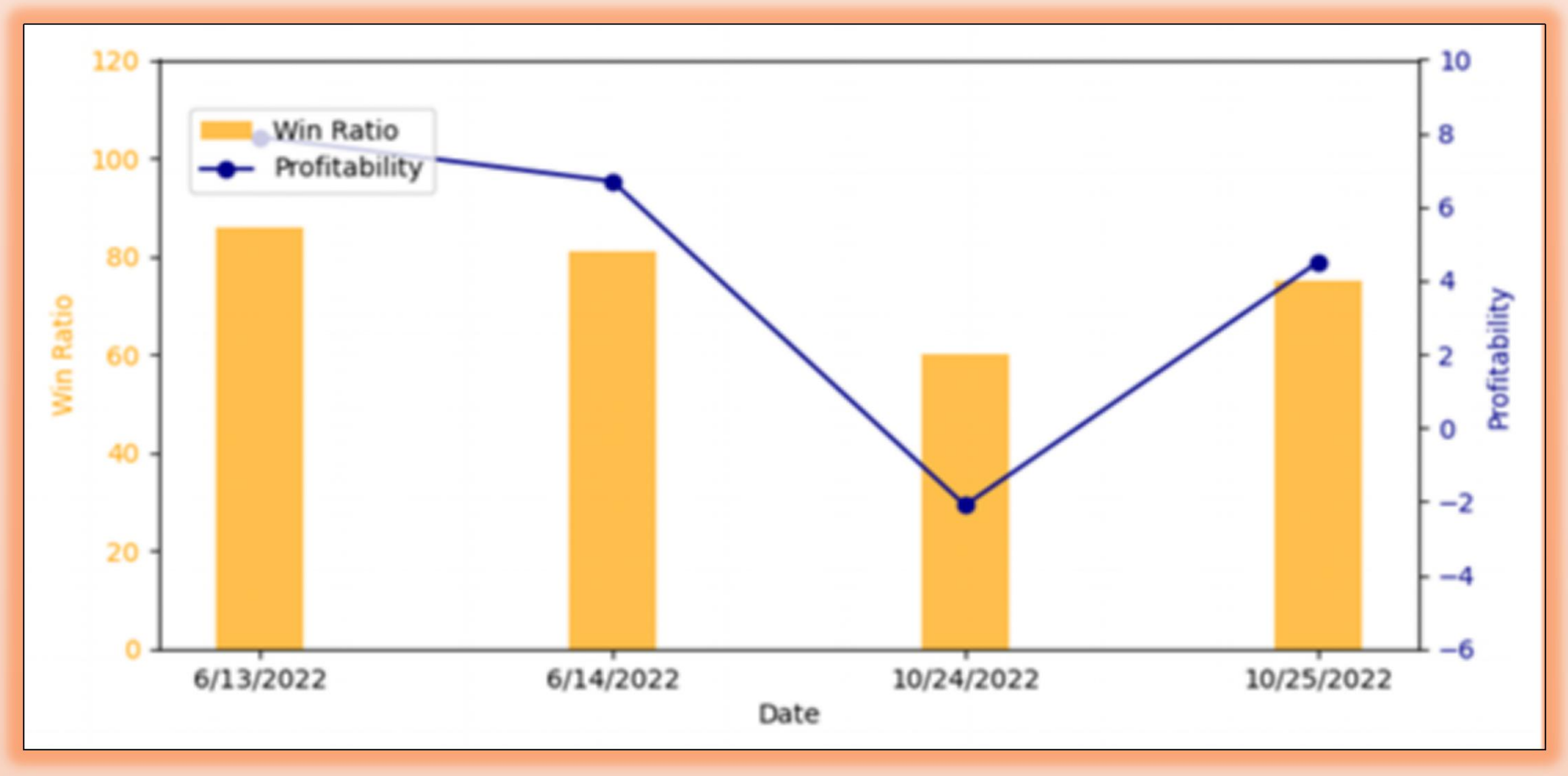
The states of Q-learning system uses market indicators such as moving averages, RSI, MACD, and volume averages to define states, capturing trends and momentum while providing the agent with a clear understanding of market conditions.

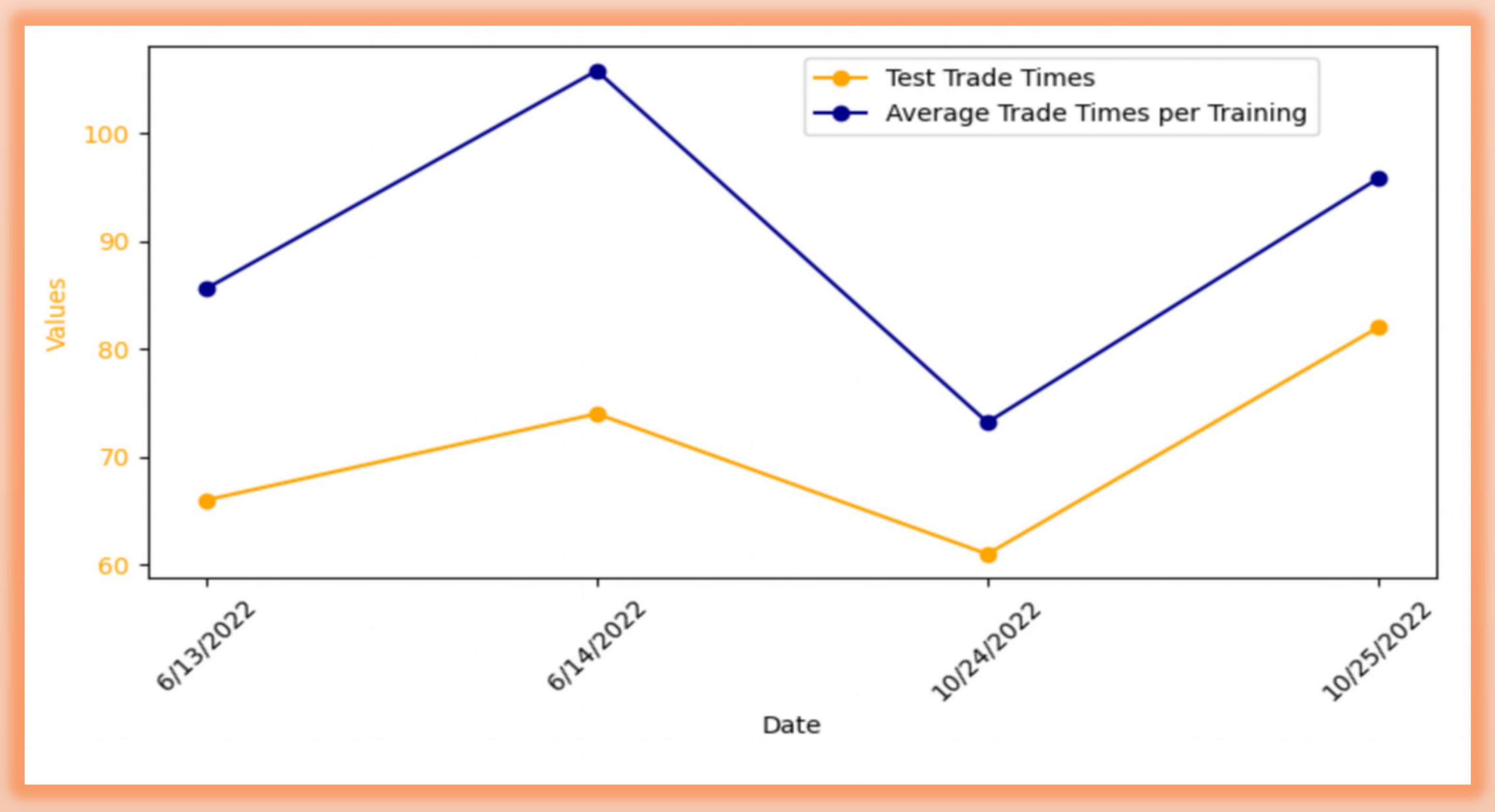
Actions include holding positions, buying on bullish signals (e.g., golden crosses), and clearing positions on bearish signals (e.g., death crosses or RSI overbought conditions). A "long-only" approach is adopted to reduce complexity and minimize risk.

The reward design incentivizes profitability and effective risk management. Positive rewards are given for gains after accounting for transaction costs, while penalties apply for losses or holding positions during unfavorable conditions. Additional rewards encourage timely profit-taking when gains exceed a target, and penalties discourage unnecessary trading.

Results

During testing, the Q-learning model achieved an 80% win rate on three days, demonstrating accuracy in trend identification and trading decisions. With 100 million initial capital, three days ended with positive returns, peaking at 8%, while the worst performance resulted in only a 2% loss.





Parameters

Key parameters are tuned for balance and efficiency.
The learning rate (α) is set to 0.01-0.05.
discount factor (γ) at 0.995.
The exploration rate (ϵ) starts at 0.01-0.1.
Transaction costs are 12.5 units per trade.
Initial capital set at 1 billion.
Moving Average Windows set at 500&10,000 ticks.
Short-term indicators using 14-20 ticks.
These settings ensure balanced profitability, risk control, and adaptability.

Conclusion

The experiment confirms that Q-learning-based trading strategies achieve high win rates and stable returns, adapting well to varying market conditions. Trades align with signals like golden crosses and RSI overbought levels. Balanced frequency and risk management ensure stability, with future improvements focusing on rewards, short-selling, and higher-frequency data.