Multi-Timeframe Algorithmic Trading Bots using Thick Data Heuristics with Deep Reinforcement Learning

Computer Science MSc Candidate: Gregory Roy

Exam committee members: Dr. Jinan Fiaidhi (Supervisor and Chair), Dr. Carlos Zerpa (External Examiner) and Dr. Sabah Mohammed (Internal Examiner)

Monday, November 28th, 2022

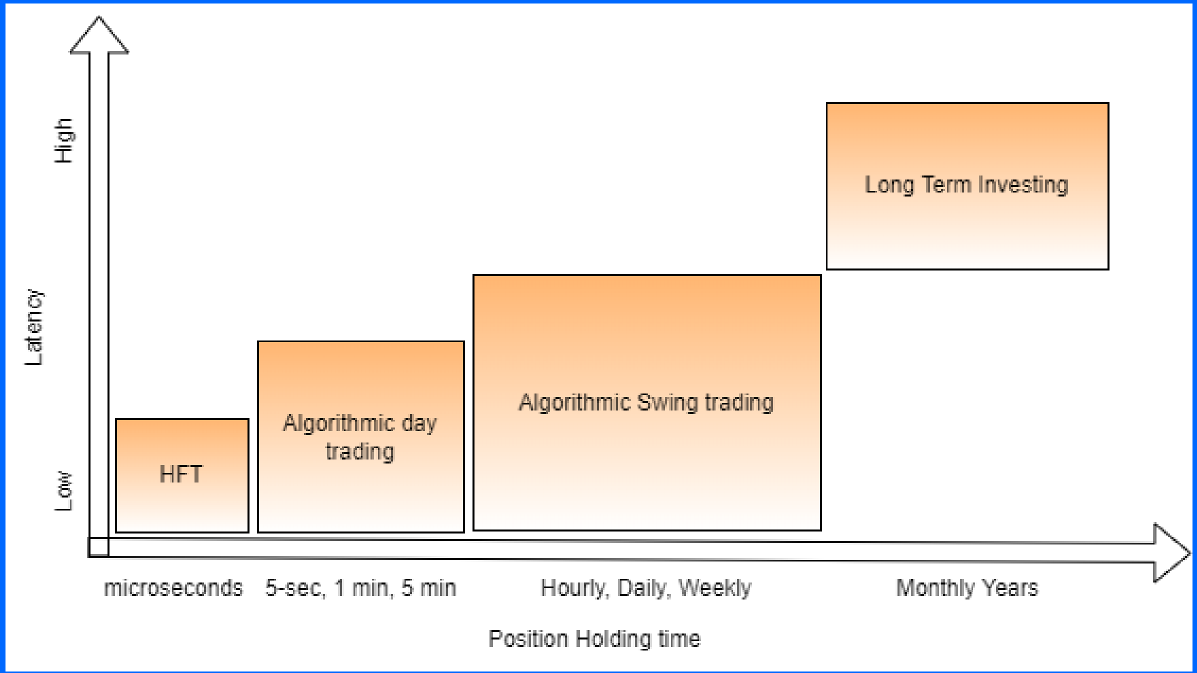
9:30 am ET

Agenda

Introduction

- Thick Data Heuristics (TDH) is a combination of human and bot-shared data management and decision-making
- Can Deep Reinforcement Learning (DRL) intraday stock trading bots and swing trading bots can be employed successfully to execute profitable trades in the NASDAQ and NYSE stock markets, on a single stock?
- Explore how adding TDH affects the performance of the trading bots.
- TDH helps add subject matter experts' experience into the decision-making process.

Latency vs Position Holding time for Day, and Swing Trading



Background

- Explore the performance of algorithmic trading bots, to see how the bots perform for day trading using intraday price movement and swing trading with weekly price bars going back ten years.
- Use Q-Learning (QL), where the goal is to find the optimal Q-value function for the output state (buy-sell-hold) with an Artificial Neural Network (ANN) in the stock market environment with iterative updates based on the Bellman equation.

TDH Combines qualitative and quantitative data.

Concept that employs heuristics and qualitative data to provide more in-depth insights and reveals hidden patterns that can be missed with quantitative techniques.

Aims at discovering the added-value heuristics to answer focused questions that can be missed by quantitative analytic techniques including ML.

Thick Data Heuristics (TDH)

Thick Data Heuristic	Agent
Multi-timeframe technical analysis	Bot
Patterns	Bot
Volume	Bot
Price action,	Bot
Price levels	Bot
Tape-reading	Bot
Fibonacci extension and retraction levels	Human
Volume profile	Human
Market structure	Human
News	Human
Sentiment	Human
Relative performance	Bot
Candlestick analysis	Human
Fundamentals	Human
Momentum	Bot

Thick Data Heuristic	Agent
Twitter	Human
Asian and European markets	Human
Index and commodity futures	Human
Hyperparameter tuning	Human
Message boards and chatrooms	Human
Volatility	Bot
Top Volume Rate Market Wide Scanner	Bot
Risk Management	Bot

Related Work for Single Stock Trading

Article	Algorithm	Dataset	Period	Time-Series interval for RL
[44]	TDQN	30 stocks	5 years	Daily
[37]	GDQN and GDPG	15 stocks	8 years	Daily
[39]	PPO, A2C, DDPG	30 stocks	7 years	Daily
[38]	TFJ, DRL	S & P 500	1999-2018	Daily
[40]	DQL	US: ETF, KOR: IDX	2010-2017	-
[34]	NN, RCNN	Nasdaq: GE, Nasdaq: GOOGL	-	-
[36]	DQL, DNN	S & P 500, KOSPI, Euro Stoxx 50, HSI	1987-2017	Daily
[47]	DQN, A3C, SDAEs, LSTM	US (AAPL, PG, IBM, ES, IF, S & P 500)	2008-2018	Minute, Daily
[39]	DRL, PPO	US (S & P 500, 5 stocks, Gold, BTC)	1960-2019	-
Proposed multi-timeframe method	TDH, DQN	Nasdaq, NYSE (7 stocks × 5 sec and 1 stock × 1 min, 10 stocks × weekly)	2012-2022	Part 1a: 5 sec, Part 1b: 1 min data, Part 2: Weekly

Novel Research:

- Combining TDH with DRL to compare the performance of different RL algorithms.
- Comparing the performance of different DRL algorithms over different timeframes.
- Determining optimal TDH-DQN trading bot parameters and settings through iterative training and testing on an idealized dataset and different market environments.

This thesis research seeks to justify the research problem to conduct this investigation by highlighting the volatility of the stock market and the limitations of AI bots to use emotional information from traders to make more accurate decisions.

Research Questions

Address the problem of how to solve the sequential decision-making problem of trading formalized as an MDP optimization of order timing executions for a single catalyst stock.

1. Can RL agents successfully learn trading order execution timing policies?
2. Which DRL algorithm performs the best across multiple timeframes?
3. Does adding TDH improve RL agent performance?

Methodology

- In Part 1, Find optimal DRL algorithm.

In Part 1a, Find the best algorithm for the five-second timeframe

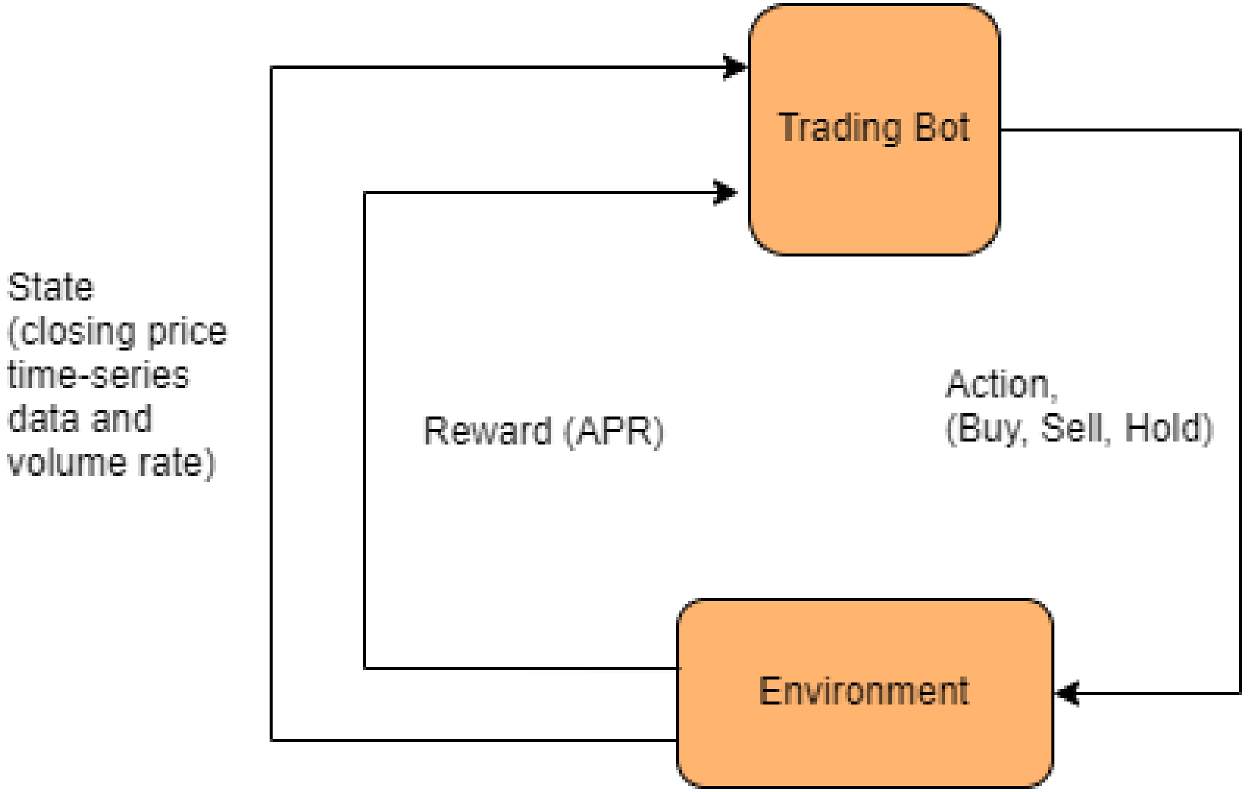
In Part 1b, Compares DRL algorithms both with and without TDH, one-minute timeframe.

- In Part 2, Compares TDH-DQN bot vs. some baseline algorithmic trading, weekly timeframe.
- In Part 3, Case Study, on user-generated ideal data

Markov Decision Process

- The foundation of RL is the Markov Decision Process (MDP).
- Markov Chain
- Markov Property
- Markov Decision Process
- State transitions depending on the last state of the system and the action of the agent.
- Given a state-action pair, the environment returns a possibly random reward to the agent.
- The agent's goal is to maximize the total reward received.

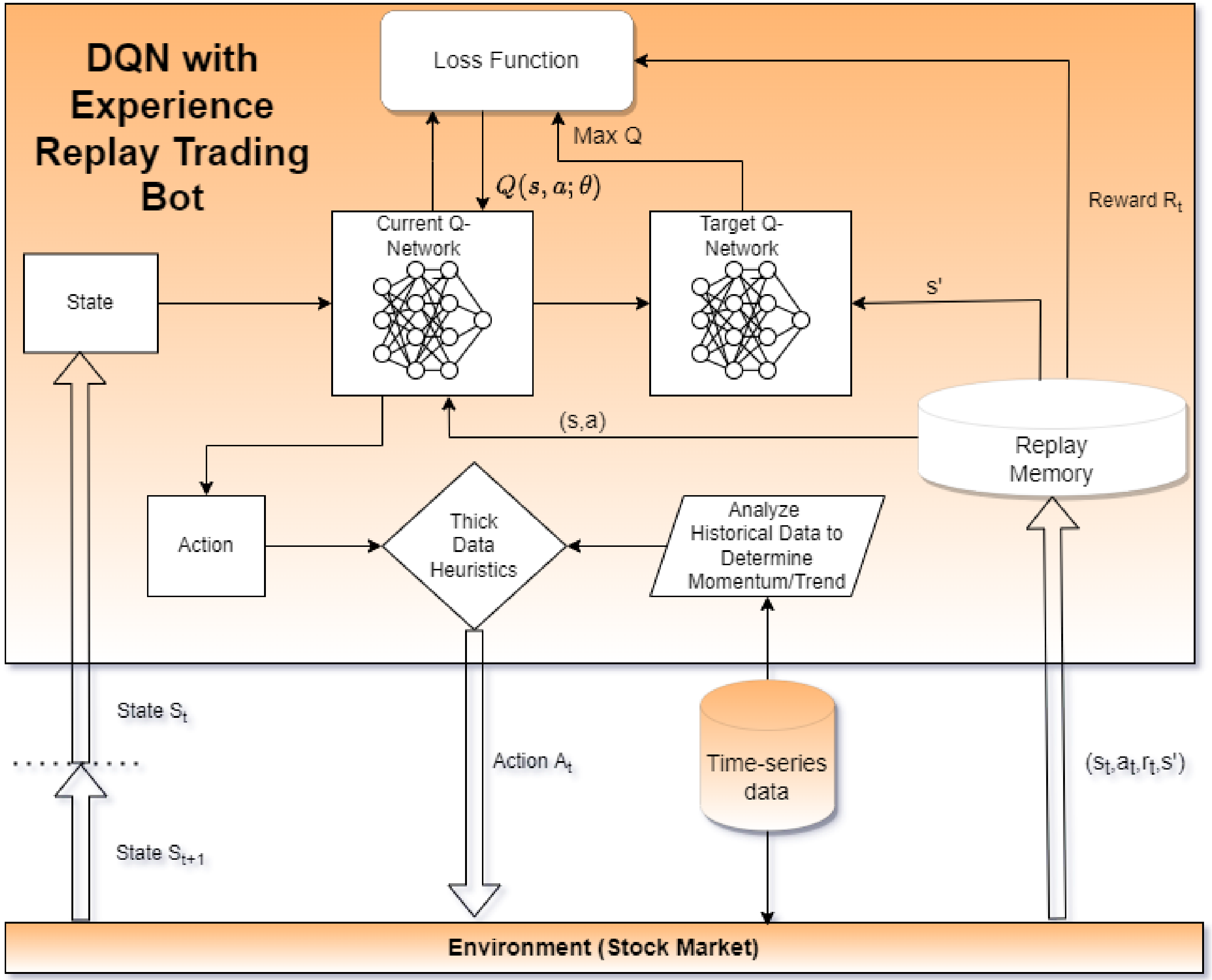
Trading Bot RL environment



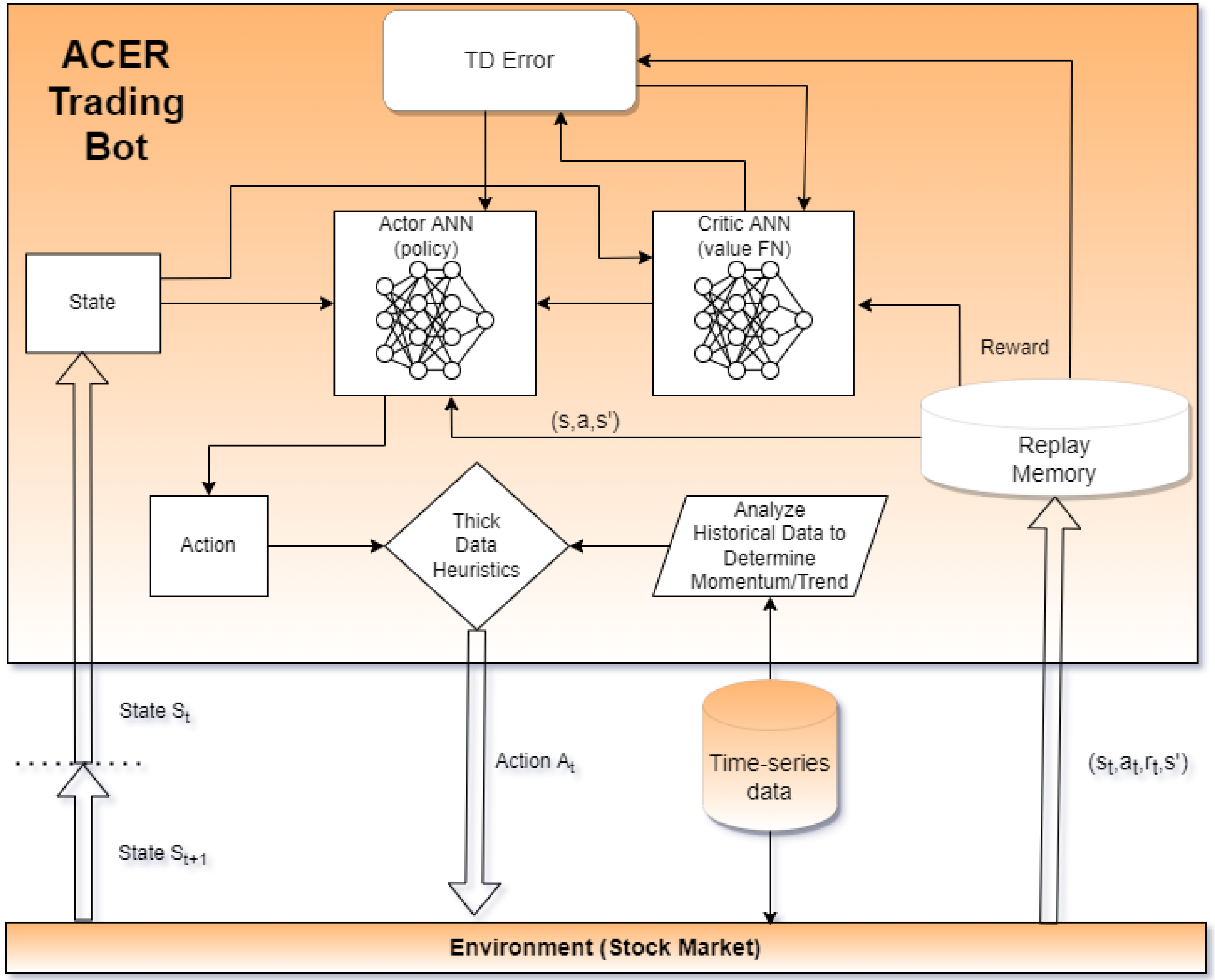
Experience Replay and Periodically Updated Q-values

- The first is Experience replay.
- The algorithm stores the history of the state, action, reward, and next-state transitions that are experienced by the agent in one large replay memory.
- Mini-batch of observations.
- Experience replay improves data efficiency, removes correlations in the observation sequences, and smooths over changes in the data distribution.
- The second method is to update Q and optimize towards target values that are only periodically updated.
- Makes the training more stable as it overcomes short-term oscillations.
- The algorithm applies gradient descent to a loss function.

DQN Trading Bot Architectures



ACER Trading Bot Architectures



Catalyst momentum stocks for Part 1

Company	Exchange	Date	Bars (5 s, 1 D, 1 W)	Trend	Catalyst
Snap Inc.	NYSE	2021-07-23	1800, 53, 28	up	Earnings beat
Alibaba Group Holding Limited	NYSE	2021-10-04	1800, 53, 29	down	China News
NIO Inc.	NYSE	2021-10-04	1800, 53, 29	down	Analyst upgrade
Advanced Micro Devices, Inc.	Nasdaq	2021-10-13	1800, 53, 29	up	New product
Plug Power Inc.	Nasdaq	2021-10-13	1800, 53, 29	up	Hot sector
SoFi Technologies, Inc.	Nasdaq	2021-10-18	1800, 53, 19	up	Analyst upgrade
FuelCell Energy, Inc.	Nasdaq	2021-10-18	1800, 53, 29	up	Hot sector

Part 1 day trading bot risk parameters

Parameter	Setting value
Holding time	Holding time = 30 minutes is the maximum holding time
First trade	5 minutes into the trading day at 9:35 am EST
Last trade	55 one-minute bars into the trading day or 10:25 am EST
Time between trades	5 minutes
Stop loss	3.0 percent
Profit target	4.0 percent

Performance metric: Accumulated Percent Returns (APR)

$$\text{Accumulated Percent Return} = [(\text{Current Price} - \text{Bought Price}) / \text{Bought Price}] * 100$$

TDH-DQN Augmented AI Pseudo Code (1 of 4)

1. Analyze the market with real-time high-volume percentage gainers and losers watchlists.
2. Human trader builds the pre-defined list of stocks.
3. Human trader performs TDH analysis with Multicharts.Net algorithmic trading platform.
4. Create policies to maximize the APR for the pre-approved catalyst stocks.
5. Load python packages.
6. Load datasets.
7. Explore data with technical analysis of multi-timeframe moving averages.

TDH-DQN Augmented AI Pseudo Code (2 of 4)

8. Data analysis to determine price levels.

9. Train-test split 70% of the dataset for training and 30% for testing.

10. DRL loop until batch complete.

11. Run the replay buffer function.

12. Qt = Updated Bellman equation.

13. Get target = Qt.

14. Calculate QB = 1, Qs = 2, QH = 0 and compare to QL tables Q-predicted and Q-target.

15. Update the Q-function by minimizing the MSE between the Q-predicted and the Q-target.

TDH-DQN Augmented AI Pseudo Code (3 of 4)

16. Fit ANN.

17. Plot buy and sell actions and total APR for each episode of the training phase.

18. Repeat until the specified number of epochs is complete.

19. Test the data.

20. Tune the model.

21. Repeat until APR > 5 percent for both training and test data.

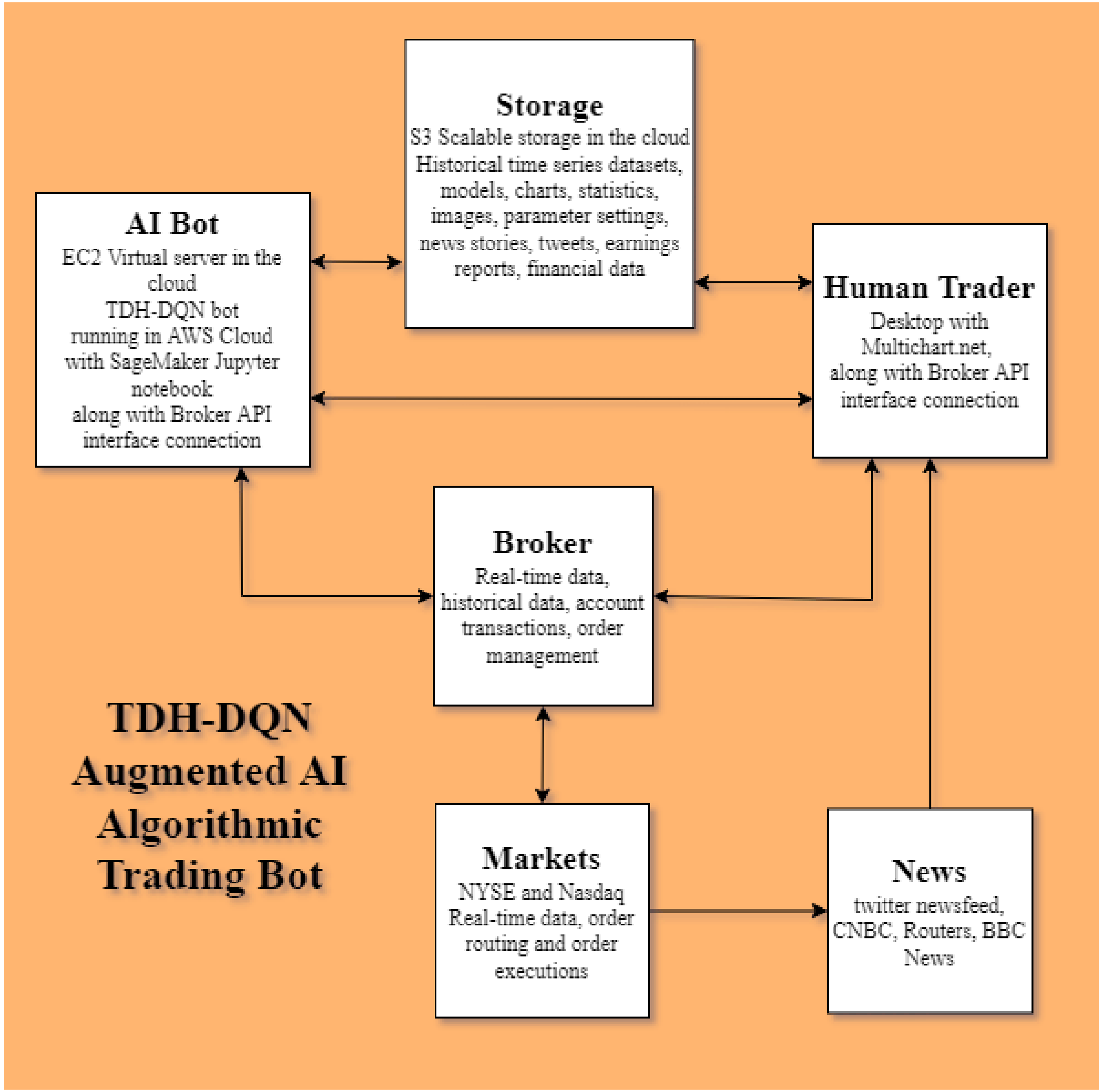
22. If the time is greater than 9:30 am EST. than execute live trades instead of testing.

23. Execute live orders when stock matches high volume rate market scanner.

TDH-DQN Augmented AI Pseudo Code (4 of 4)

24. Analyze real-time market scanners and compare them to a pre-defined watchlist.
25. Manage risk.
26. Use TDH hard-coded trading rules to filter the trade execution orders.
27. Repeat downloading of new data and backtesting for parameter optimization.
28. Train the model and execute new orders.
29. Repeat.

TDH-DQN Augmented AI bot Network Architecture



Results: TDH + DRL algorithm comparison for Part 1a 5-sec day trading bots

Trading bot algorithm	Abbreviation
Policy Gradient	PG
Q-learning	DQN
Double Q-Learning	DQL
Actor-Critic	AC
Curiosity Q-Learning	CQN
Recurrent curiosity Q-Learning	RCQL
Dual Actor-Critic	DAC
Dual Curiosity Q-Learning	DCQL
Buy and Hold	B & H
Sell and Hold	S & H

Stock	PG	DQN	DQL	AC	CQN	RCQL	DAC	DCQL	B&H	S&H
SNAP	1.36	2.09	1.39	1.54	1.09	1.43	1.05	1.27	0.44	-0
BABA	0.54	1.08	1.82	1.21	1.78	1.54	1.94	1.79	-2	1.99
NIO	2.9	2.89	2.6	2.55	2.53	2.57	2.51	2.54	-4	4
AMD	1.96	2.11	2.26	2.31	2.3	1.29	2.13	2.26	2.53	-2.5
PLUG	3.25	3.75	2.81	3.24	3.86	2.63	2.68	2.84	4.37	-4.4
SOFI	4.9	5.54	4.85	4.56	5.69	5.29	5.49	5.44	6.47	-6.5
FCEL	3.88	1.04	4.71	4.39	4.58	3.19	3.55	3.64	11.9	-12
AVG.	2.68	2.64	2.92	2.83	3.12	2.56	2.76	2.83	2.82	-2.8

Results: Part 1a 5-sec bot computing time comparison, total time between orders (measured in seconds)

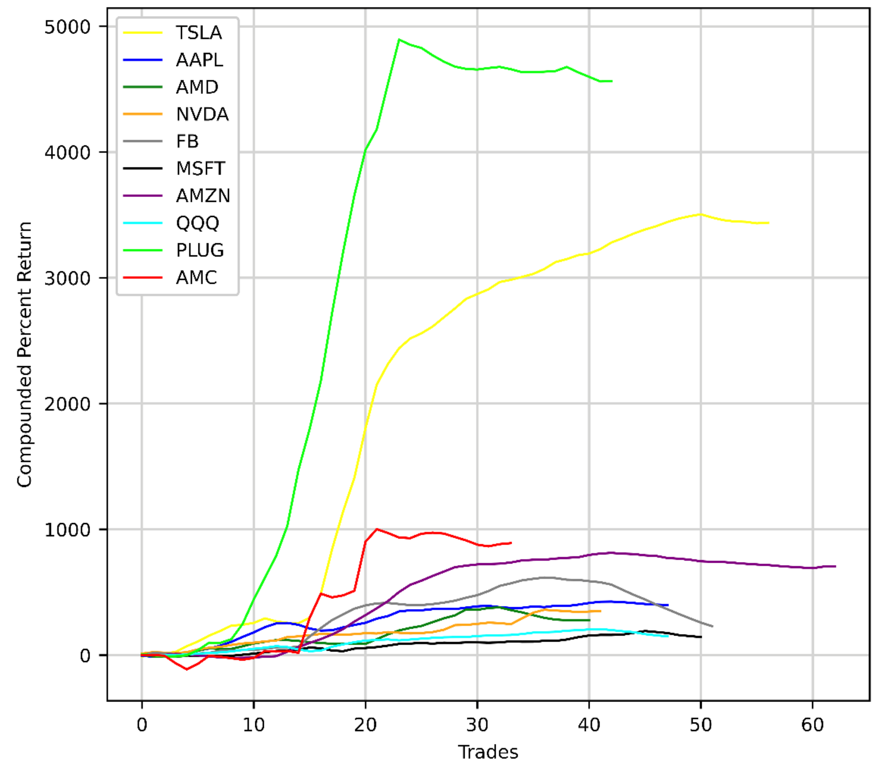
Stock	PG	DQN	DQL	AC	CQN	RCQL	DAC	DQ
SNAP	5.17	27.02	28.54	45.04	14.04	76.06	51.83	15.22
BABA	5.16	34.43	34.26	46.85	13.95	81.15	56.35	15.95
NIO	5.35	32.67	38.4	44.78	13.49	78.55	51.27	15.01
AMD	5.4	27.98	29.07	44.8	14.13	78.23	51.68	15.24
PLUG	5.46	28.2	30.2	46.46	14.24	82.54	53.67	15.25
SOFI	5.5	28.14	32.65	46.24	13.78	80.53	53.65	15.54
FCEL	5.42	27.54	30.65	45.92	13.62	80.08	51.44	15.23
Average	5.3	29.4	32.0	45.7	13.9	79.6	52.8	15.3

Results: DRL algorithm APR comparison both with and without TDH applied for Part 1b

DRL parameters	SNAP TDH	SNAP	BABA TDH	BABA	NIO TDH	NIO	AMD TDH	AMD	PLUG TDH	PLUG	SOFI TDH	SOFI	FCEL TDH	FCEL
TDH applied	Y	N	Y	N	Y	N	Y	N	Y	N	Y	N	Y	N
DQN	-1.04	1.03	0.59	-1.31	0.44	0.34	-0.1	0.69	-0.25	0.67	2.02	0.03	4.11	2.32
Duel DQN	-1.5	-4.36	0.14	0.32	0.73	-0.78	1.4	1.75	0.11	-0.57	2.4	2.36	4.34	-1.41
Recurrent DQN	-0.65	0	0.59	0	0.67	0	0.17	0	0.35	0	1.63	0	-1.52	0
Double DQN	-4.5	-0.85	0.58	-1.32	0.8	-3.63	2.1	1.62	2.4	0.81	3.08	1.56	6.09	3.84
Double recurrent DQN	-3	-1.78	0.24	-0.81	0.5	-1.66	0.91	0.76	0.97	2.44	2.47	1.56	3.79	3
Double duel DQN	-2.09	4.08	0.25	-0.38	0.53	-0.31	1.47	-0.01	-0.85	-1.59	2.05	0.8	4.1	3.86
Double duel recurrent DQN	-0.73	0	0.64	0	0.73	0	0.02	0	0.35	0	0	0	3.64	0
Curiosity DQN	-1.04	-2.51	0.64	-0.65	0.53	-1.41	1.64	1.46	0.55	-1.34	1.84	1.08	2.8	5.25
Recurrent curiosity DQN	-0.4	0.44	0.63	-0.08	0.8	-0.34	1.64	1.07	0.25	-0.6	1.53	1.08	3.79	4.95
Duel curiosity DQN	-3.35	1.59	0.35	-0.7	0.76	-0.8	0.37	1.42	1.23	0.91	2.12	-0.07	3.52	1.13
AC	-0.4	1.19	0.32	-0.56	0.86	-0.53	1.69	1.27	0.25	-0.46	1.77	2.5	3.96	2.26
Dual AC	-1.1	2.72	0.35	-0.22	0.59	-1.24	1.4	1.32	1.64	-0.93	1.81	1.49	3.97	2.93
Recurrent AC	-1.87	-1.43	0.24	-0.23	0.76	-0.97	0.91	0.19	-0.25	-2.33	1.49	2.16	3.14	-2.87
Duel recurrent AC	-0.73	0.24	0.16	-0.75	0.53	-2.1	1.51	1.24	1.53	1.32	1.18	1.95	2.96	2.86
Avg.	-1.60	0.03	0.41	-0.48	0.66	-0.96	1.08	0.91	0.59	-0.12	1.81	1.18	3.48	2.01
THD applied Average =	0.92													
No THD applied Average =	0.37													

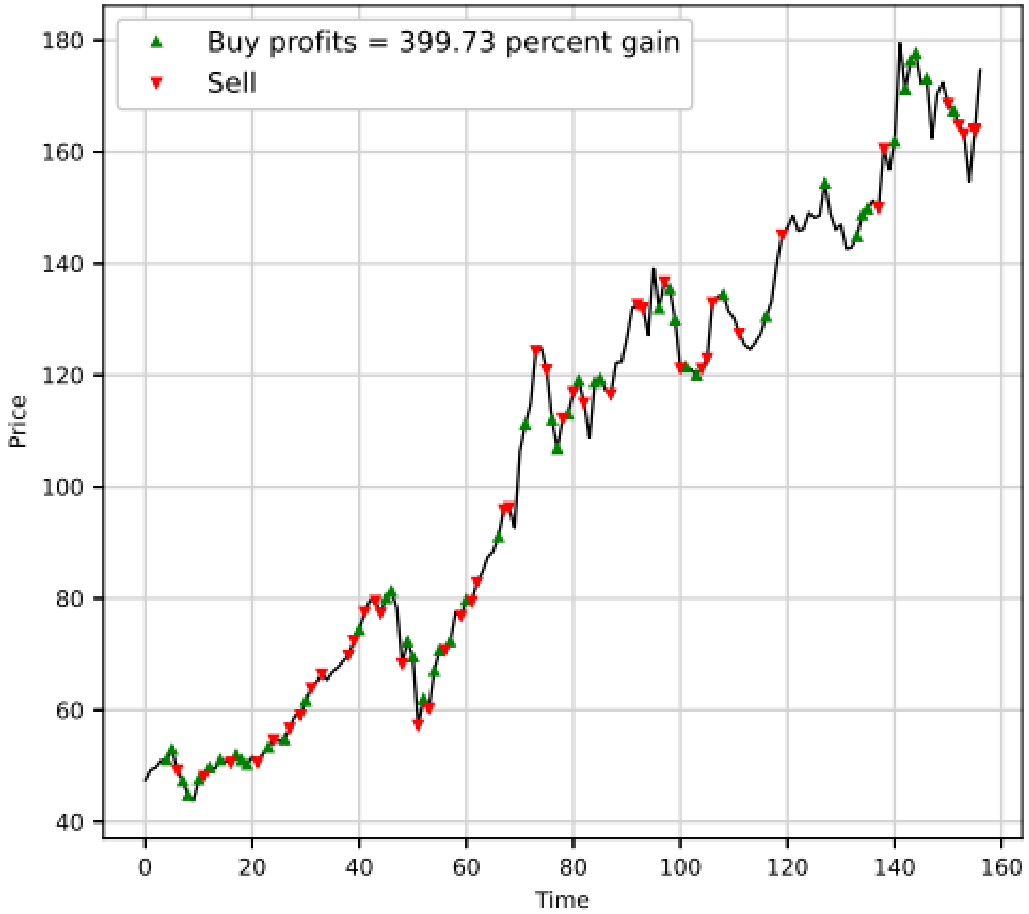
Results: Part 2 APR comparison of TDH-DQN bot vs. benchmark strategies

Stock	B & H	TDH-DQN	Pivot reversal (2)	MACD LE	Outside bar
TSLA	819	3,439	911	814	1,163
AAPL	181	400	861	452	864
AMD	161	278	809	296	208
NVDA	335	350	1,208	765	712
FB	39	231	141	-52	-60
MSFT	100	144	388	265	310
AMZN	73	707	30	12	101
QQQ	93	150	215	-28	175
PLUG	609	4,564	11,769	1,386	554
AMC	464	891	819	640	1,148
Average	287	1,115	1,715	455	518



Results: CPR vs. trades for Part 2

Results: Part 2 orders



Part 2 AMD Price vs. Time Chart: TDH-DQN Swing Trading Bot's result

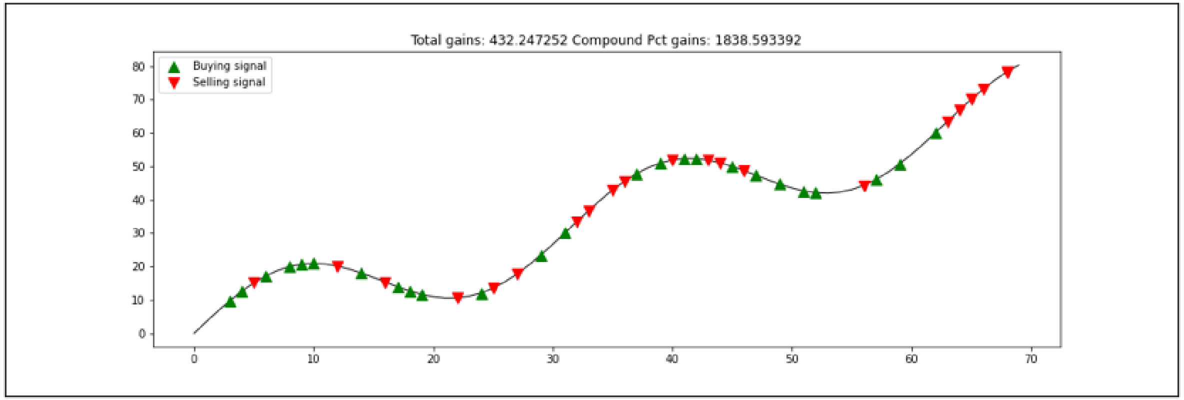


Figure 11. Part 3 Orders: Episode 1 of TDH-DQN Training.

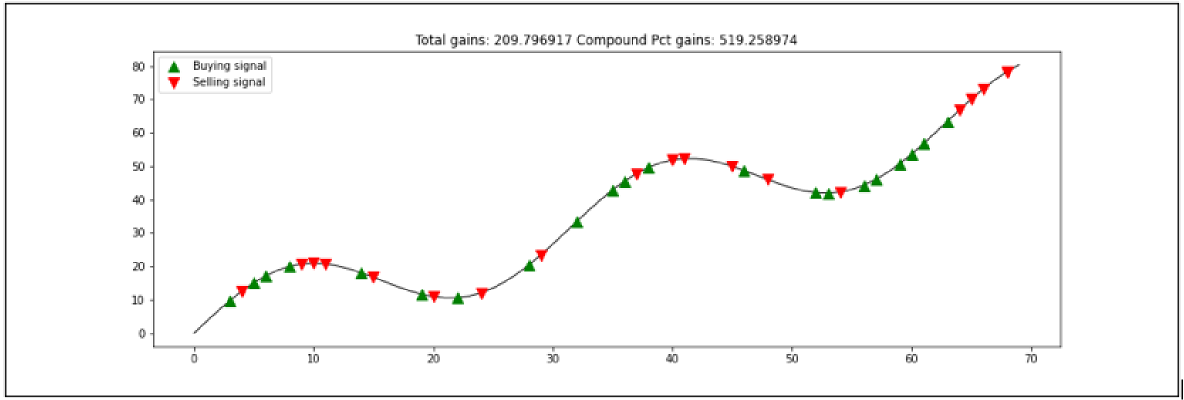
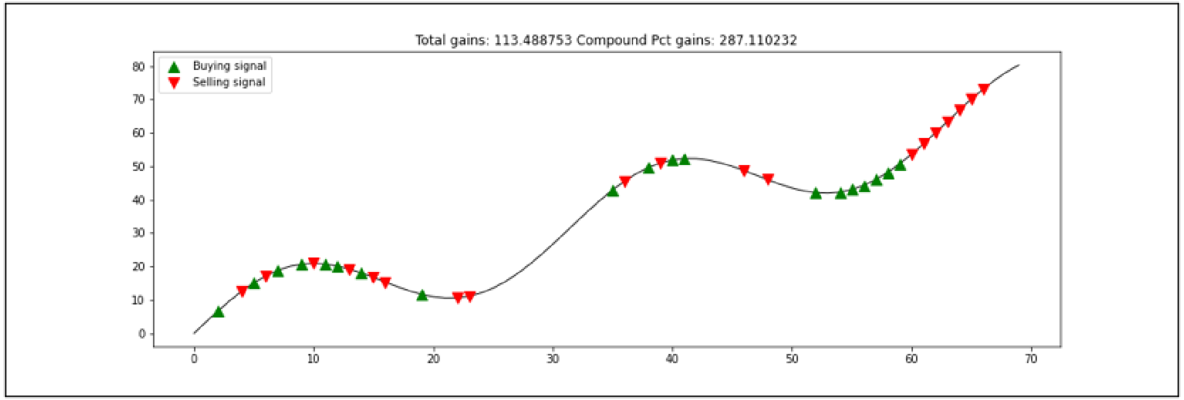


Figure 12. Part 3 Orders: Episode 2 of TDH-DQN Training.



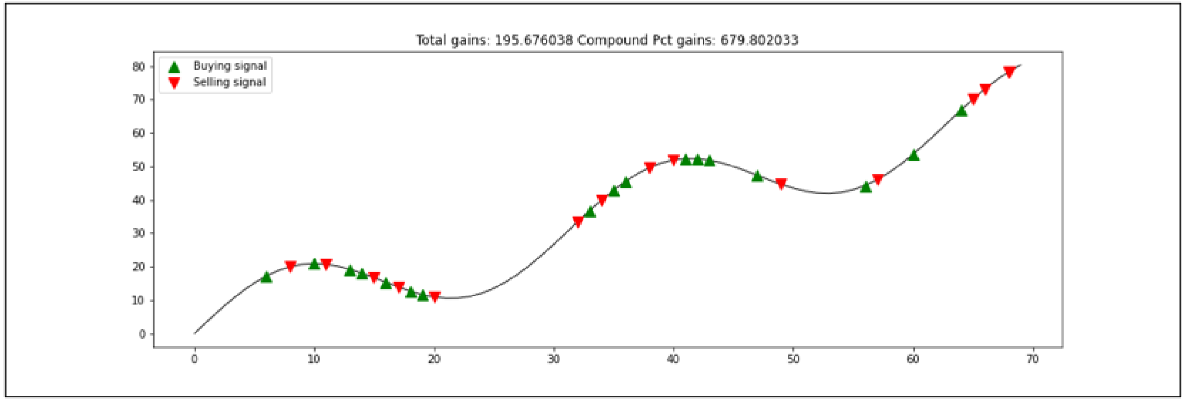


Figure 14. Part 3 Orders: Episode 4 of TDH-DQN Training.

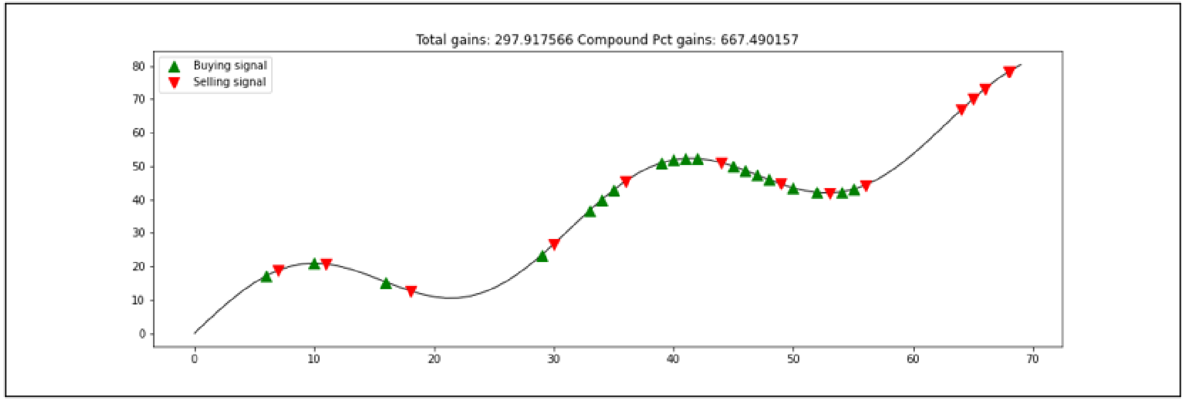
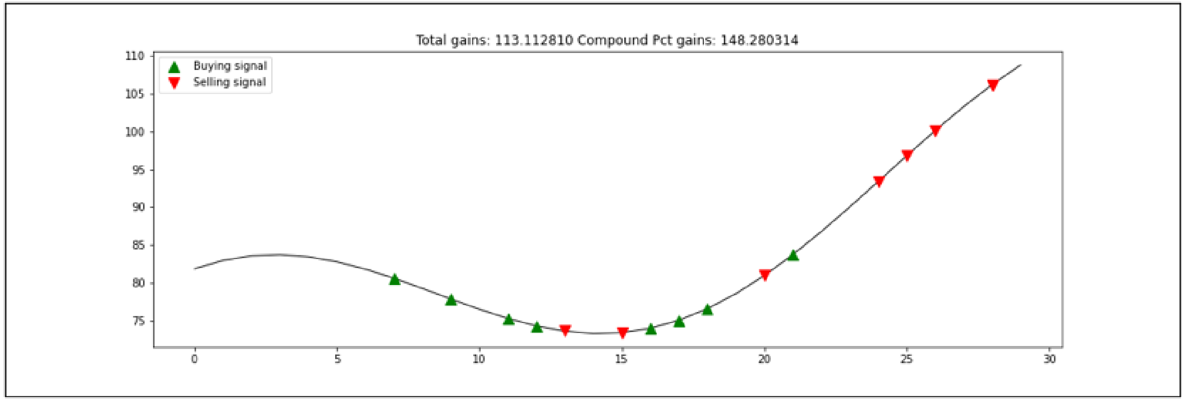
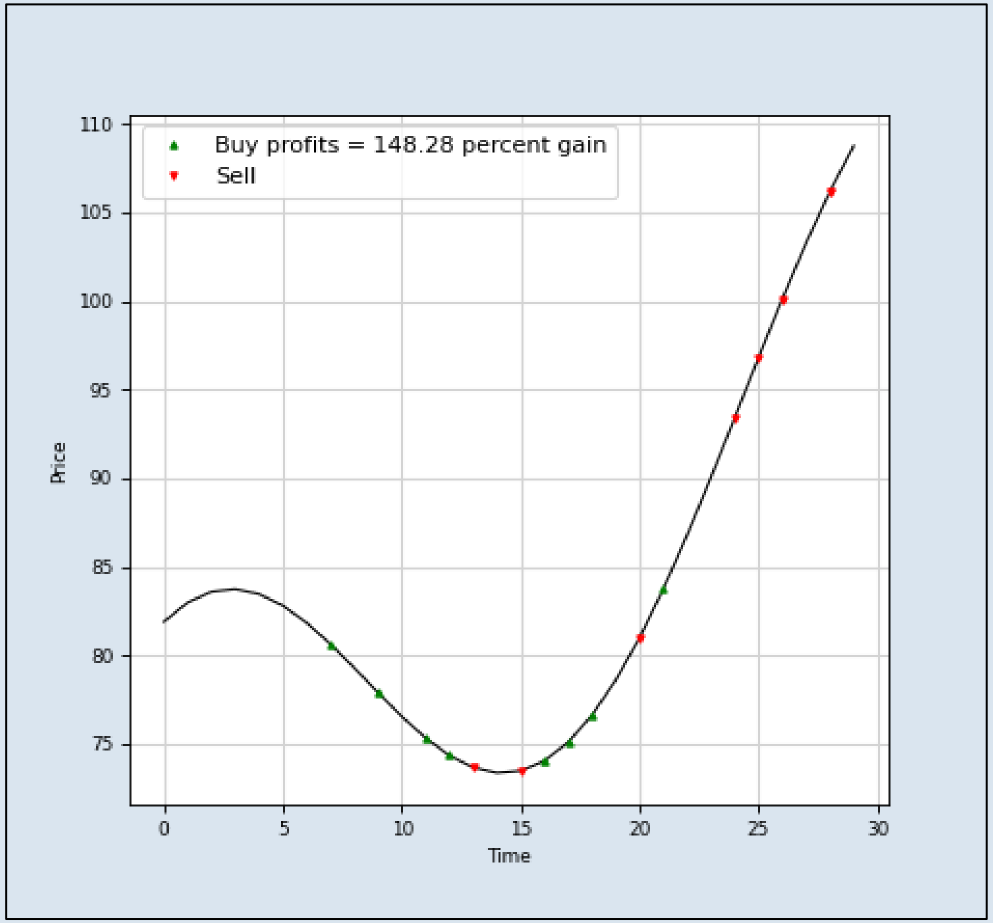


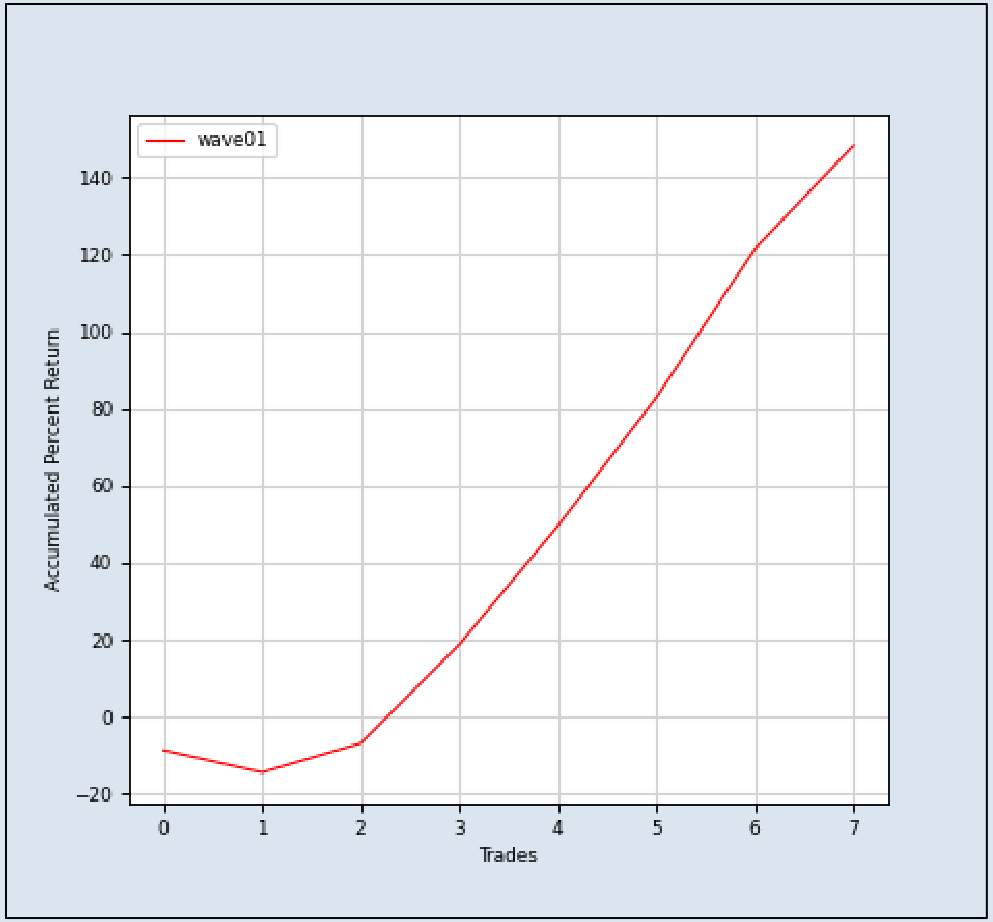
Figure 15. Part 3 Orders: Episode 5 of TDH-DQN Training.



Results: Part 3 Training orders: Bots Training Episodes

Results: CPR vs. Trades for Part 3 Case Study Test Data





Conclusions

- RL works where the data is few, and the behavior is complex.
- Catalyst stocks behave differently
- There is limited data in the morning when the stock opens
- Must adapt to the limited data

Conclusions

- Learn policies that are more complex and powerful than what a human trader could learn.
- Without intuition, it’s difficult for trading bots to learn the intuitive relationships between input and their corresponding output.
- Deep neural networks negatively compound a model’s explainability further.
- Overfitting isn’t something the bots can overcome on their own.
- Intuition and heuristics software capabilities are limited.
- An ML algorithm is only as good as the data it is trained on.
- Major issue in trading known as the dreaded overfitting, to be avoided at all costs.

Conclusions

- Training data used is important.
- The DRL bots memorize and overfit the patterns in the data, especially if you overtrain them.
- It's important to train the bots on the type of data the market is trading currently
- The bots must be able to recognize market regime changes.
- The bot must determine when the short-term direction is aligned with the long-term direction for momentum.
- The DRL bots are good at extracting new knowledge from the input data.

Insights

- Domain expertise.
- Choices of model objectives and performance.
- Predictive performance continues to improve with more data.
- Model and data complexity need to match.
- Managing data quality
- Human traders rely on emotions.
- Human traders struggle with following all the rules.
- Algorithmic DRL trading bots excel at many rules-based aspects of trading.
- Experience, history, and making mistakes is often the only path to success
- DRL bots are a valuable extension or augmentations of human intelligence.

Summary

- Cannot replace human intuition and heuristics.
- Complex models, if not correctly guided can over-fit
- Crafting financial models is an art form more than a science.
- It’s important to test many different model parameters.
- There is no perfect system in trading that works all the time.
- Human intervention and domain expertise.
- Define objectives, select, and curate data, design and optimize a model, and make appropriate use of the results.
- Avoid over-fitting when the signal-to-noise ratio is low with financial data.
- Patterns evolve quickly as signals decay.

Future work

- Systematization and automation of the remaining responsibilities assigned to the human agent in the TDH Table.
- Further interests include position sizing, cloud-based AI, CNNs for visual candlestick analysis, and sentiment analysis using Natural Language Processing (NLP).
- Include AI bots capable of trading multiple stocks in a portfolio at the same time.

Future work

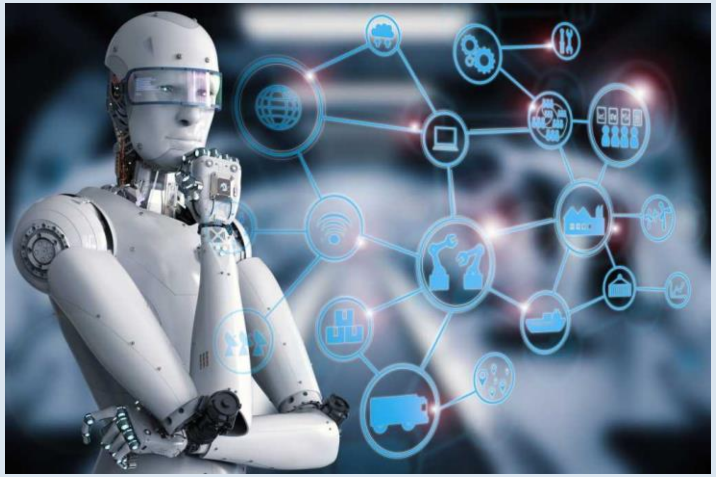
- Including measures of motivation and emotions of traders in RL algorithms using physiological measures could allow the bots to achieve more informed decisions.
- If the human and AI bot traders were operating in a **Virtual Reality (VR)** environment, eye-sensing technology could be used to track the humans.
- Future research could explore the comparison of heart rate measures to thick data heuristics and deep reinforcement learning as primary inputs when training AI bots.
- Future research using physiological measures could make it possible for AI bots to make more informed execution timing decisions.



Published Article

Roy G, Fiaidhi J, Mohammed S. Multi-Timeframe Algorithmic Trading Bots Using Thick Data Heuristics with Deep Reinforcement Learning. Artificial Intelligence Evolution. 2022 Nov 12:107-59.

https://ojs.wiserpub.com/index.php/AIE/article/view/1722



ACKNOWLEDGEMENTS

- Thesis Supervisors:

[Dr. Jinan Fiaidhi] and

Dr. Sabah Mohammed

- Family:
- External Reviewer:

Dr. Carlos Zerpa

Thank You

Gregory Roy

gjroy@lakeheadu.ca