

Reinforcement Learning for Trading

Team:

Mariem Ayadi, Shreyas S. Jadhav, Benjamin W. Livingston,
Amogh Mishra & Kevin Womack

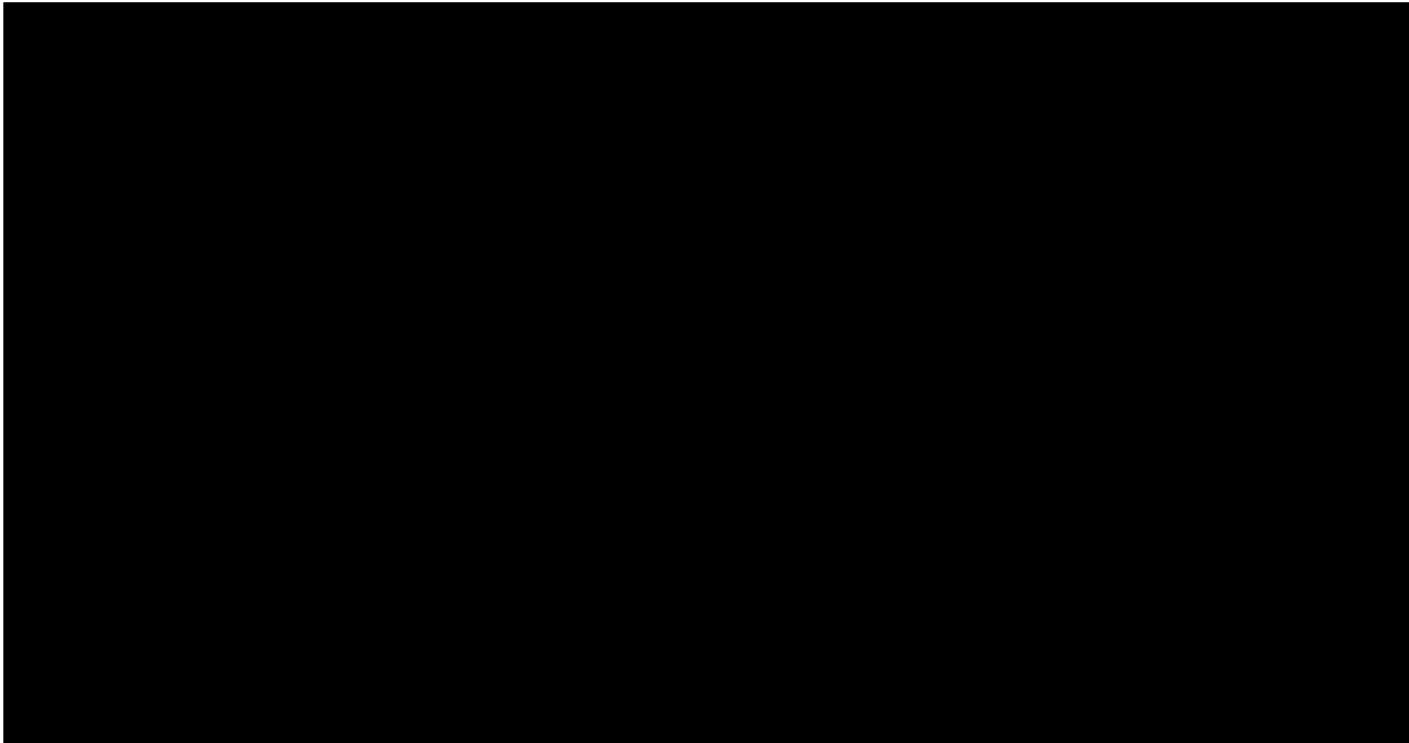
Industry Mentors:

Naftali Cohen, Srijan Sood & Zhen Zeng

DSI Supervisor:

Adam Kelleher

Fall 2020



Message from our Industry Mentors

Algorithmic Trading

~~Uncontrollable~~



~~Opaque~~

Value Industry

RL for Trading

Problem, Background and Impact

- How would you teach a 10 year old to trade stocks?
 - How can we transfer this logic to artificial intelligence?
- Smart trading is nuanced but situational



Technical Details

RL Algorithm: Q-Table

	0	1	2
states			
000	6.535896e-15	1.150069e-15	9.502829e-15
001	4.821914e-15	8.724745e-15	2.123327e-15
002	4.070962e-16	3.971945e-15	2.331322e-15
003	8.417407e-15	2.070823e-15	7.424695e-15
004	3.921541e-15	1.822565e-15	7.435394e-15
...
945	2.127106e-15	1.063860e-15	7.645156e-15
946	6.745562e-16	5.844151e-15	8.128164e-15
947	4.760910e-15	2.664735e-16	1.211542e-15
948	6.710698e-15	4.610479e-15	8.495217e-15
949	9.820431e-15	2.627686e-15	9.903721e-15

Technical Details

Q-Table: Actions

	Hold	Buy	Sell
states	0	1	2
000	6.535896e-15	1.150069e-15	9.502829e-15
001	4.821914e-15	8.724745e-15	2.123327e-15
002	4.070962e-16	3.971945e-15	2.331322e-15
003	8.417407e-15	2.070823e-15	7.424695e-15
004	3.921541e-15	1.822565e-15	7.435394e-15
...
945	2.127106e-15	1.063860e-15	7.645156e-15
946	6.745562e-16	5.844151e-15	8.128164e-15
947	4.760910e-15	2.664735e-16	1.211542e-15
948	6.710698e-15	4.610479e-15	8.495217e-15
949	9.820431e-15	2.627686e-15	9.903721e-15

Technical Details

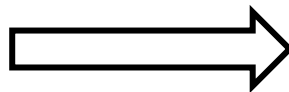
Q-Table: States

	0	1	2
states			
000	6.535896e-15	1.150069e-15	9.502829e-15
001	4.821914e-15	8.724745e-15	2.123327e-15
002	4.070962e-16	3.971945e-15	2.331322e-15
003	8.417407e-15	2.070823e-15	7.424695e-15
004	3.921541e-15	1.822565e-15	7.435394e-15
...
945	2.127106e-15	1.063860e-15	7.645156e-15
946	6.745562e-16	5.844151e-15	8.128164e-15
947	4.760910e-15	2.664735e-16	1.211542e-15
948	6.710698e-15	4.610479e-15	8.495217e-15
949	9.820431e-15	2.627686e-15	9.903721e-15

Technical Details

Q-Table: q-values

At Initialization



After Training

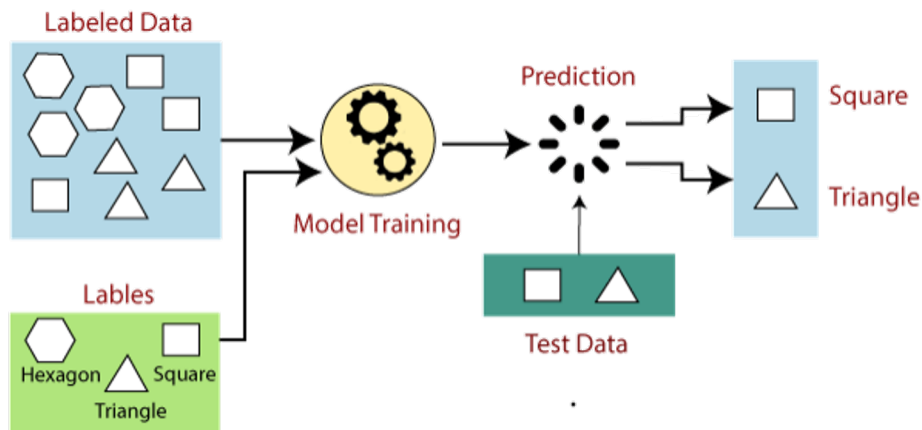
	0	1	2
states			
000	6.535896e-15	1.150069e-15	9.502829e-15
001	4.821914e-15	8.724745e-15	2.123327e-15
002	4.070962e-16	3.971945e-15	2.331322e-15
003	8.417407e-15	2.070823e-15	7.424695e-15
004	3.921541e-15	1.822565e-15	7.435394e-15
...
945	2.127106e-15	1.063860e-15	7.645156e-15
946	6.745562e-16	5.844151e-15	8.128164e-15
947	4.760910e-15	2.664735e-16	1.211542e-15
948	6.710698e-15	4.610479e-15	8.495217e-15
949	9.820431e-15	2.627686e-15	9.903721e-15

	0	1	2
states			
125	1039.152295	-199.416058	0.000000
836	818.478203	-17.252940	0.000000
012	0.000000	0.000000	772.035554
015	114.642682	0.000000	482.541923
820	0.000000	813.994734	-256.499334
826	0.000000	522.391103	0.000000
520	-429.199782	306.093788	563.782309
220	2261.389249	-98.369717	-1752.117028
225	485.628902	0.000000	-99.410019
421	344.160373	0.000000	0.000000

What is RL in the first place?

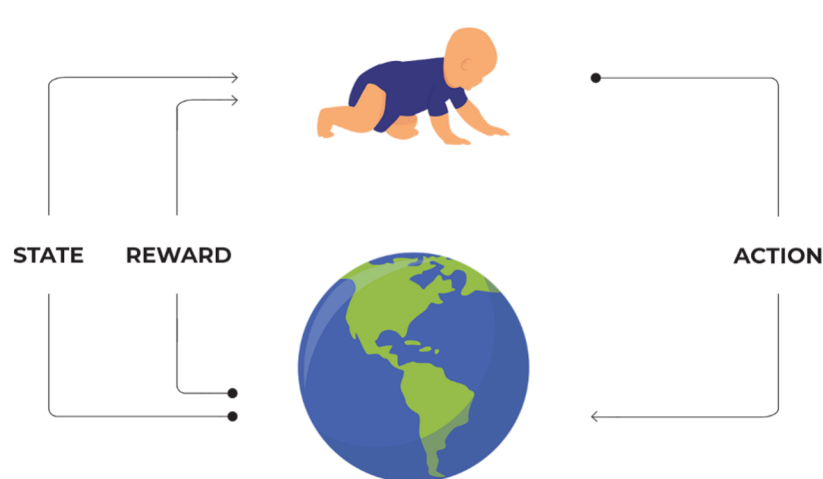
Comparison with Machine Learning

Machine Learning: Classification



Source: Javatpoint.com

Reinforcement Learning (RL)



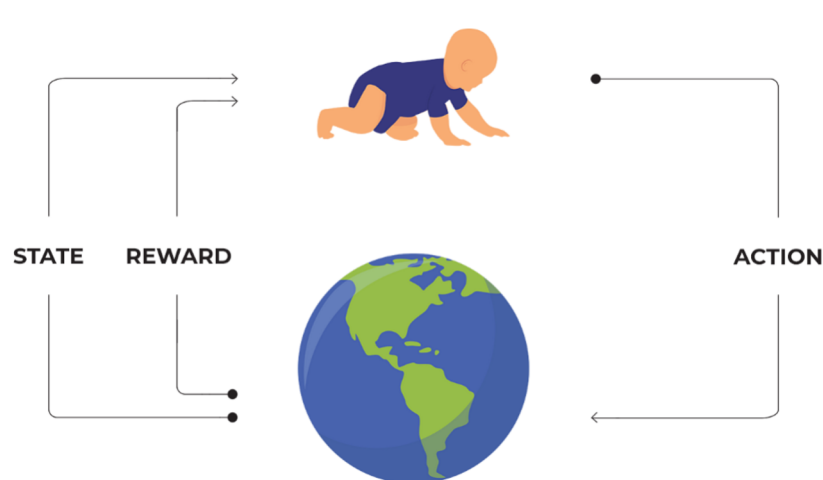
Source: Alexey Poddiachyi - towardsdatascience

What is RL in the first place?

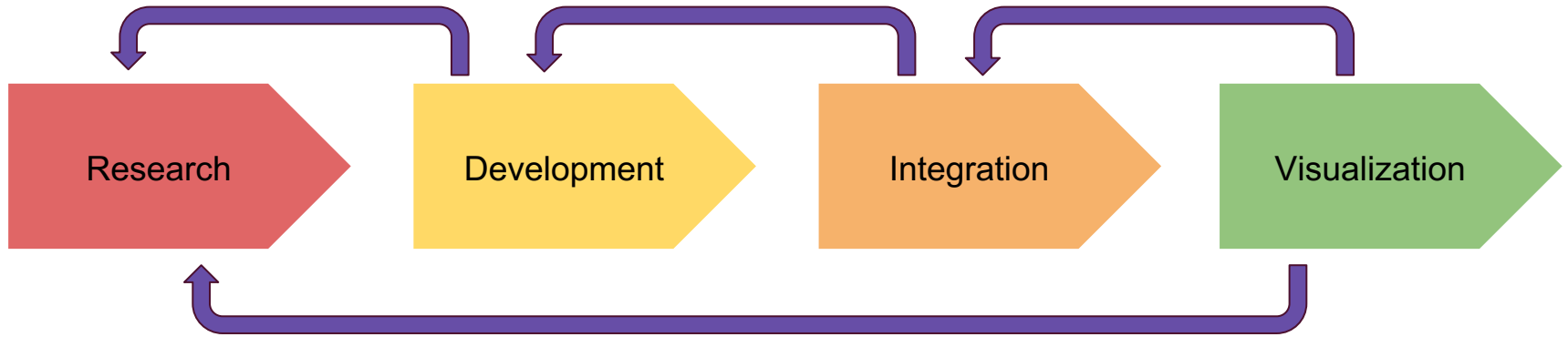
Comparison with Machine Learning

- What is unique about RL?
- How does it related to Q-tables presented?

Reinforcement Learning (RL)



Source: Alexey Poddiachyi - towardsdatascience



- What technical indicators should be used for our state representation?
- What signals would we expect certain metrics to give an agent?

- What baselines should we construct to compare performance?
- How should the notion of reward be defined?

- What information is needed from the Q-learner to examine its policy?
- What structure is best for code review and further development?

- Is the Q-Learning agent doing something meaningful/logical?
- What actions were taken by the agent?

Demo

Statistical Results

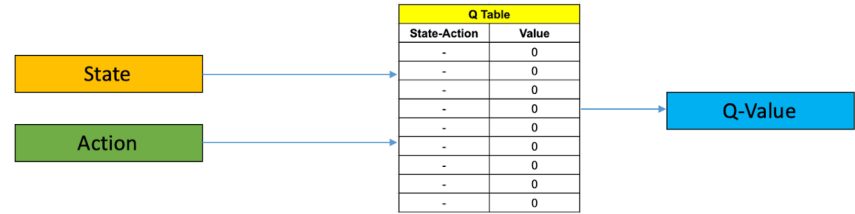
Q-Learner Vs Benchmarks

	Hold Consistently	Random Action	Rule-Based	OLS	Q-Learner
Buy %	0%	34%	51%	53%	43%
Sell %	0%	34%	48%	47%	33%
Hold %	100%	34%	48%	47%	24%
Sharpe Ratio	+0.73	+0.74	+0.77	+0.89	+1.11
Information Ratio	-0.34	-0.32	-0.24	+0.07	+0.52
Mean Daily Return	+0.043%	+0.044%	+0.045%	+0.052%	+0.064%
Mean Return Days After Buying	N\A	+0.059%	+0.049%	+0.050%	+0.045%
Mean Return Days After Holding	+0.043%	+0.013%	-0.0204%	-0.123%	+0.023%
Mean Return Days After Selling	N\A	+0.058%	+0.046%	+0.056%	+0.117%
Volatility	0.007	0.007	0.007	0.009	0.01
T-test	>0.27	>0.24	>0.25	>0.3	>0.32
Levene Test	<0.141	<0.51	<0.94	>*0.01	>*0.01

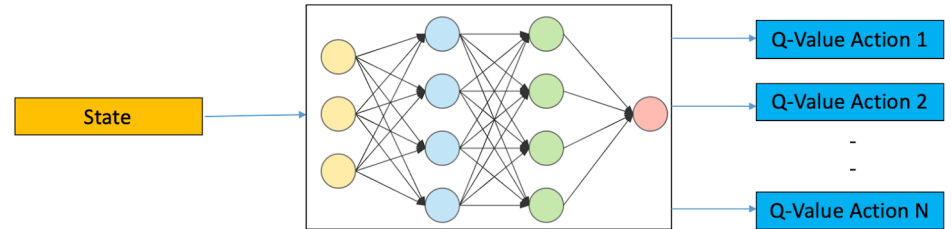
Challenges & Future Work

RL Algorithm: DQN

- Q-learning + Deep Learning = DQN
- But why do we need DQN?



Q Learning



Deep Q Learning

Challenges worth tackling

If we had more time....

- This is just one stock - just the beginning!
 - Managing a portfolio
- Further tuning of hyperparameters
- Extending state & action space
 - Sentiment analysis
 - New indicators
 - Buy/Sell in different amounts
 - Etc...

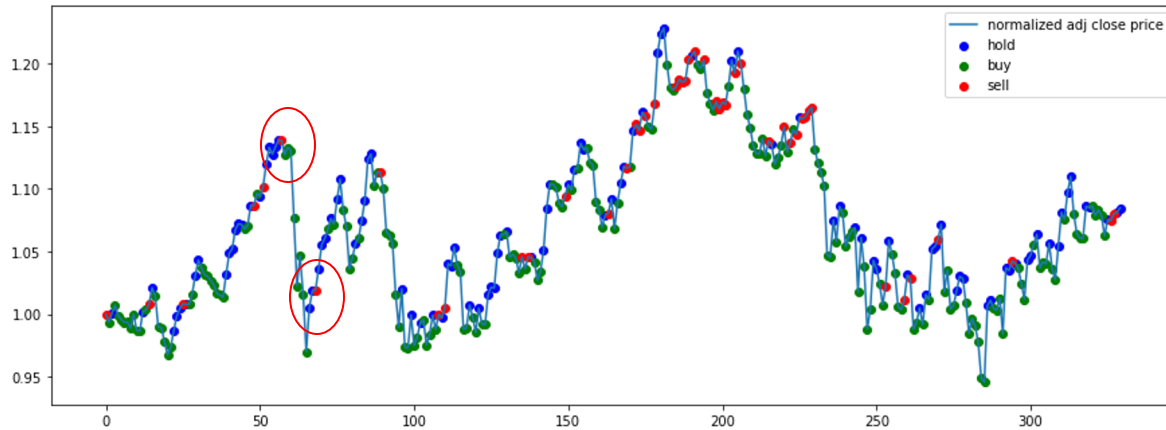
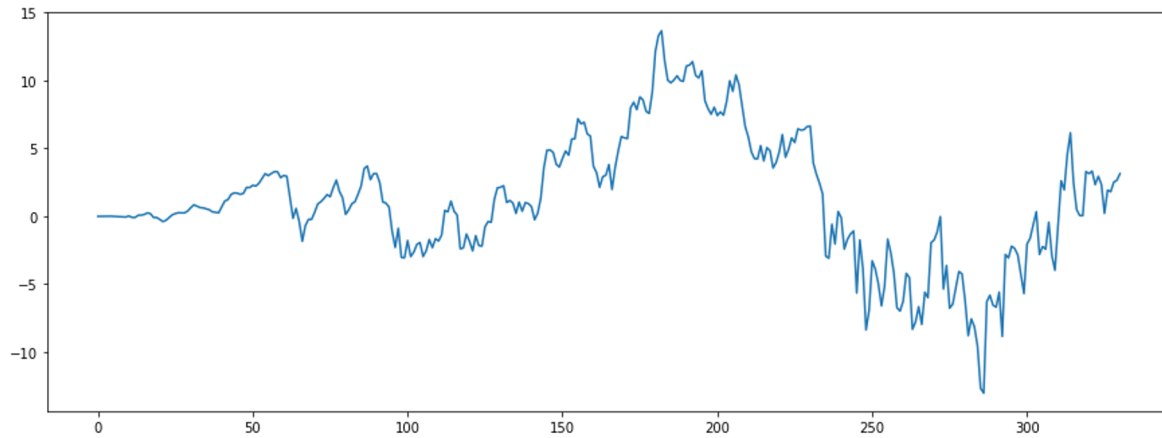
Appendix

Goal

- Build an RL agent that maximizes cumulative return on a portfolio with a single stock/asset investment (ex: JPMC, AAPL, GOOG, FB)
- Visualize the policy learned by the agent

Environment:

- Initial cash input $\$100,000$
- In-sample period: 01/01/2007 – 12/31/2016
- Out-of-sample (testing) period: 01/01/2017 – 12/31/2019
- Allowable actions: Buy, Sell, Hold



Source: <https://medium.com/@nyxqian/stock-trader-with-q-learning-91e70161762b>