Schulich
School of Business

FINE 6900 W19: Individual Study

Application of Machine Learning to Portfolio Optimization

March 29, 2019

Daniel Fudge

215868904

Executive Summary

Big data, compute power and algorithmic advances have driven an explosion of research and applications in machine learning. The purpose of this paper is to gain a deeper understanding of how these new machine learning techniques can be applied to financial portfolio management.

The structure of this paper provides a background and introduction of the different types of machine learning and then focuses on reinforcement learning, which is perfectly suited for portfolio management. Once the required machine learning concepts are reviewed, the paper surveys current applications in the finance industry and ends with the next steps in this research.

Table of Contents

Background

Machine Learning (ML) is a subset of Artificial Intelligence (AI) that learns from data without being explicitly programed. A McKinsey paper stated that companies invested $26B to $39B in AI in 2016 and noted a growing gap between early adopters and others, which poses an urgent challenge to those not investing in AI.[1] The figure below illustrates the related exponential revenue growth projected by a Tractica research report.[2] These numbers emphasize the need for governments, industry and academia to understand these emerging trends and how they could disrupt their environment.

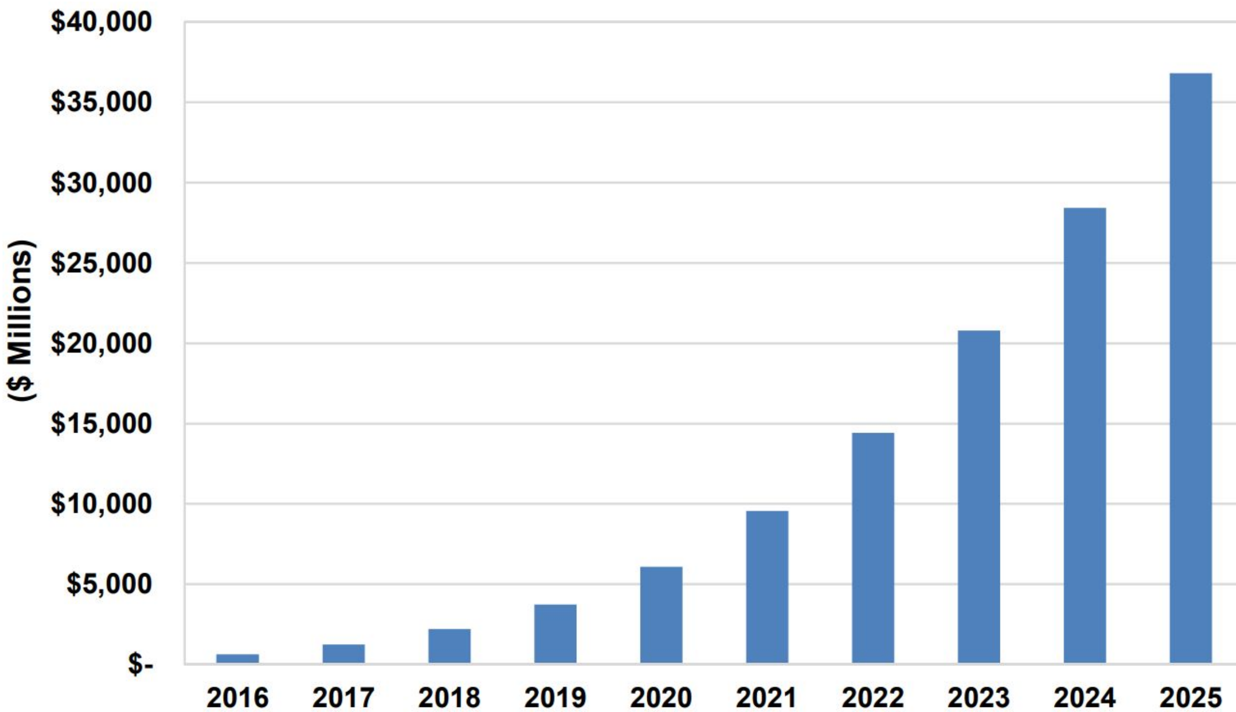


Figure 1. Projected world AI related revenues.
Image source: https://www.tractica.com/wp-content/uploads/2016/08/MD-AIMF-3Q16-Executive-Summary.pdf

[1]Jacques Bughun at el. "Artificial Intelligence The Next Digital Frontier?" *McKinsey Global Institute*, Discussion Paper, June 2017.
[2]Kaul, Aditya and Clint Wheelock. "Artificial Intelligence Market Forecasts." *Tractica Research Report*. 2016.

The Tractica report investigated 191 use cases of AI across 27 industries and identified the top 10 AI use cases in terms of projected 2025 revenue. Interestingly algorithmic trading was identified as the top use case. This clearly reinforces the need for those of us interested in portfolio management to understand its potential applications or risk being left behind.

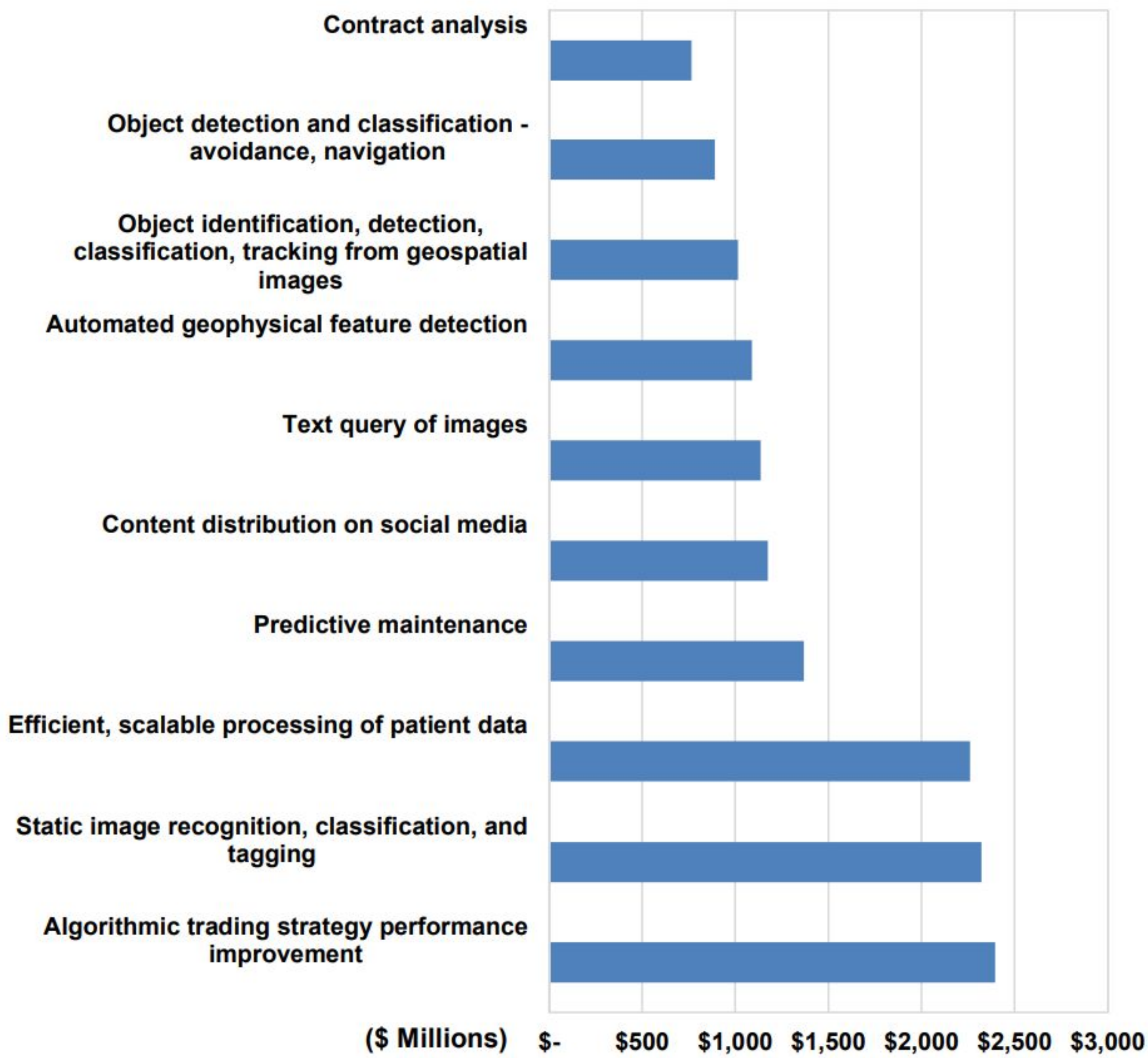


Figure 2. Top 10 AI use cases based on projected 2025 revenue.
Image source: https://www.tractica.com/wp-content/uploads/2016/08/MD-AIMF-3Q16-Executive-Summary.pdf

These rapidly increasing levels of investment and projected revenue are the result of an explosion of ML capabilities and applications during the last decade. While there are many articles discussing why AI and ML are experiencing such unprecedented growth, it comes down to three converging forces: big data, compute power, and algorithmic advances.[3,4,5]

ML is only as good as the data you feed it and we are in the middle of a data deluge. 90% of the world's data has been generated in the last two years alone and the data generated by U.S. companies is enough to fill 10,000 Libraries of Congress.[6] Paralleling this flood of data is the infrastructure to efficiently store and access it.

This data alone would be unmanageable and useless with not only the raw increase in compute power from Moore's law but also the rise of GPUs and Cloud computing, which is commoditizing high performance supercomputing capabilities. Now anyone can have access to a high-end big data pipeline with zero investment costs. Figure 3 below shows how the 19% cloud computing growth dominates the 3% average IT growth.[7]

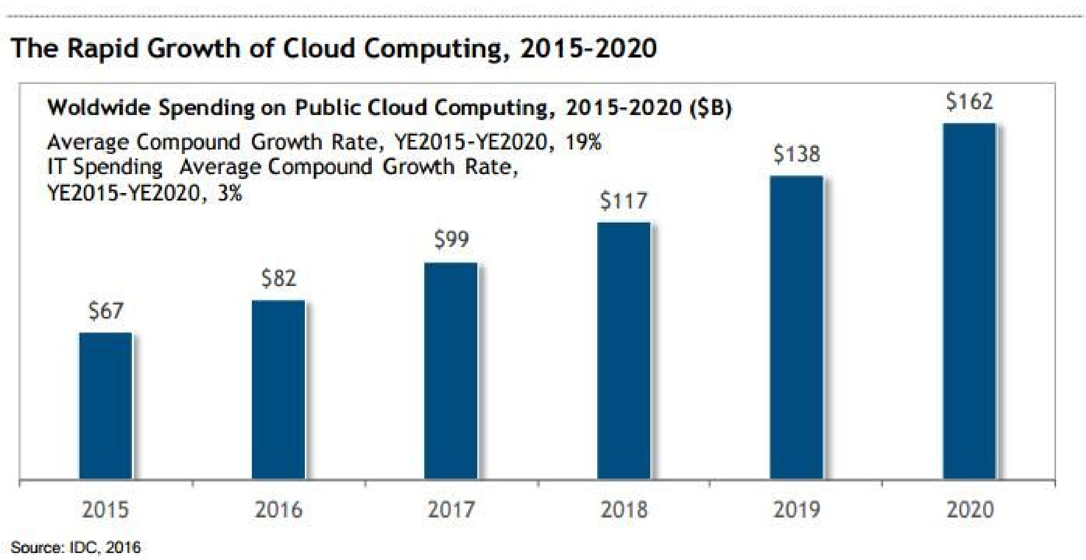


Figure 3. Growth of cloud computing.
Image source: http://blogs.forbes.com/louiscolumbus/files/2017/04/growth-of-cloud-computing.jpg

This rapid expansion of data and easily accessible compute capacity sparked a renaissance in AI research. Many of the pioneering ideas such a Artificial Neural Networks (ANNs) have histories reaching back as far as the 1940's but were

3 Janakiram MSV. "In The Era Of Artificial Intelligence, GPUs Are The New CPUs." *Forbes*, August 7 2017.
4 Rodrigo Beceiro. "What is Artificial Intelligence and why now?" *Medium*, August 8 2018.
5 Babak Hodjat. "The AI Resurgence: Why Now?" *Wired*, March 2015.
6 Carole Gunst. "10 Eye-opening Stats About the Growth of Big Data." *Attunity*, August 20, 2018.
7 Louis Columbus. "Roundup Of Cloud Computing Forecasts, 2017." *Forbes*, April 29, 2017.

abandoned in the 70s and 80s because they were impractical at the time.[8] Now we have both the data and compute power to apply them to real problems. With this revival, the last 5 years has seen a near doubling of research papers as shown in figure 4.[9] The open source community has also helped accelerate this development, especially Google AI's release of Tensorflow in 2015 and nearly 5000 publications and counting.[10]

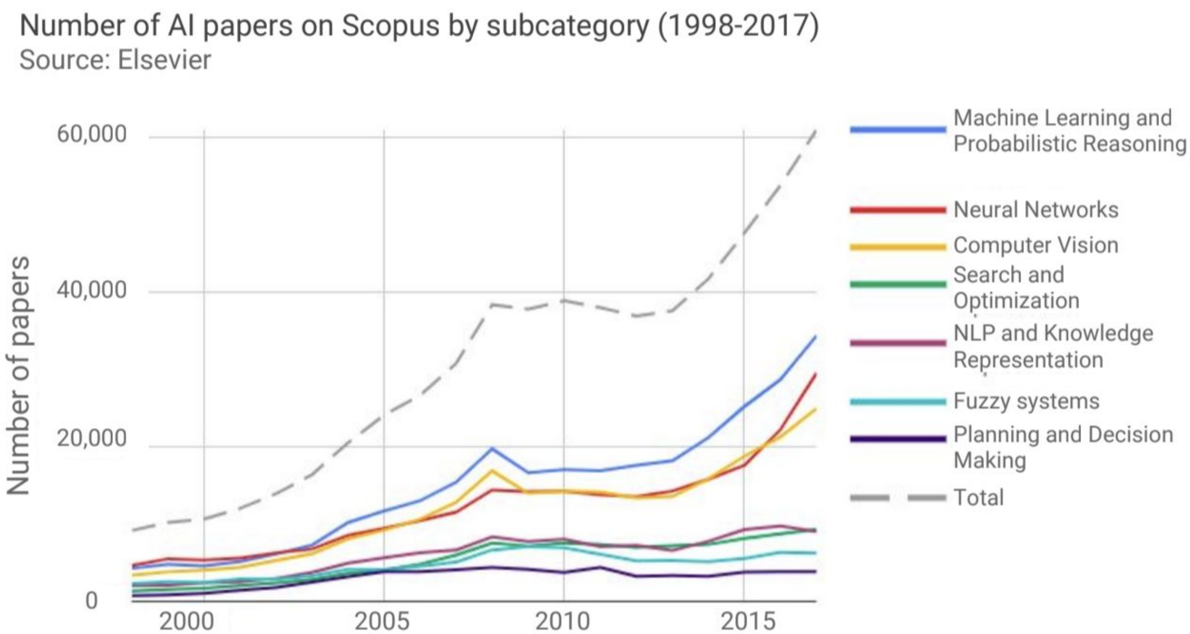


Figure 4. Growth of AI research.
Image source:
https://www.theverge.com/2018/12/12/18136929/artificial-intelligence-ai-index-report-2018-machine-learning-global-progress-research

[8] McCulloch, Warren and Walter Pitts. "A Logical Calculus of Ideas Immanent in Nervous Activity." *Bulletin of Mathematical Biophysics*. Vol. 5, Issue 4 (1943): 115–133. doi:10.1007/BF02478259

[9] Statt, Nick. "The AI boom is happening all over the world, and it's accelerating quickly." *The Verge*, December 12, 2018.

[10] Google AI. "Publication database." Accessed march 3, 2017. https://ai.google/research/pubs/

Motivation

Clearly we are witnessing the emergence of a new powerful capability that is poised to continue to grow in the coming years and alter how we work, play and possibly even see the world. The world of finance, especially portfolio management, is built upon data. The immense complexity, nonlinearity and opacity of the interactions that drive markets is often considered so overwhelming that it appears random. So the question becomes can the emergent machine learning capabilities that are purpose built for immense and highly dimensional datasets be applied to portfolio management.

Unfortunately, this question in not reflected in the current MBA and concurrent graduate diploma in financial engineering course content. Therefore, the motivation of this research is to gain a deeper understanding of how machine learning can be applied to financial portfolio management.

Much of the understanding used as a basis for this this report was obtained from the three online courses listed below and detailed in the appendix. This was also supplemented by the many resources listed in the Bibliography.

1. MIT Sloan & CSAIL: Artificial Intelligence: Implications for Business Strategy
2. AWS: Developer Associate Certification
3. NYU Tandon: Machine Learning and Reinforcement Learning in Finance

ML Fundamentals

Before discussing how ML can be applied to portfolio management, we must first review some fundamentals and terminology starting with AI. There is a joke that says "if it's written in Python it's ML, if it's written in PowerPoint it's AI." This is obviously an exaggeration but hits on the nebulous nature of the term AI. In the Massachusetts Institute of Technology (MIT) AI course, it referred to AI as a "suitcase word". Marvin Minsky coined the term "suitcase word" to describe words into which people pack many meanings.[11] Chethan Kumar's article "Artificial Intelligence: Definition, Types, Examples, Technologies" discusses many of these definitions and my favorite due to its simplicity is "the capability of a machine to imitate intelligent human behavior."[12]

The figure below illustrates one possible hierarchy of AI systems. This image is helpful to obtain a general exposure to different AI systems, but it is in no way complete. For instance, the ML branch is missing semi-supervised and reinforcement learning.

[11] Rob Campbell. "Defensiveness: A suitcase word." *Mawer*, July 18 2018.

[12] Chethan Kumar. "Artificial Intelligence: Definition, Types, Examples, Technologies." *Medium*, August 2018.

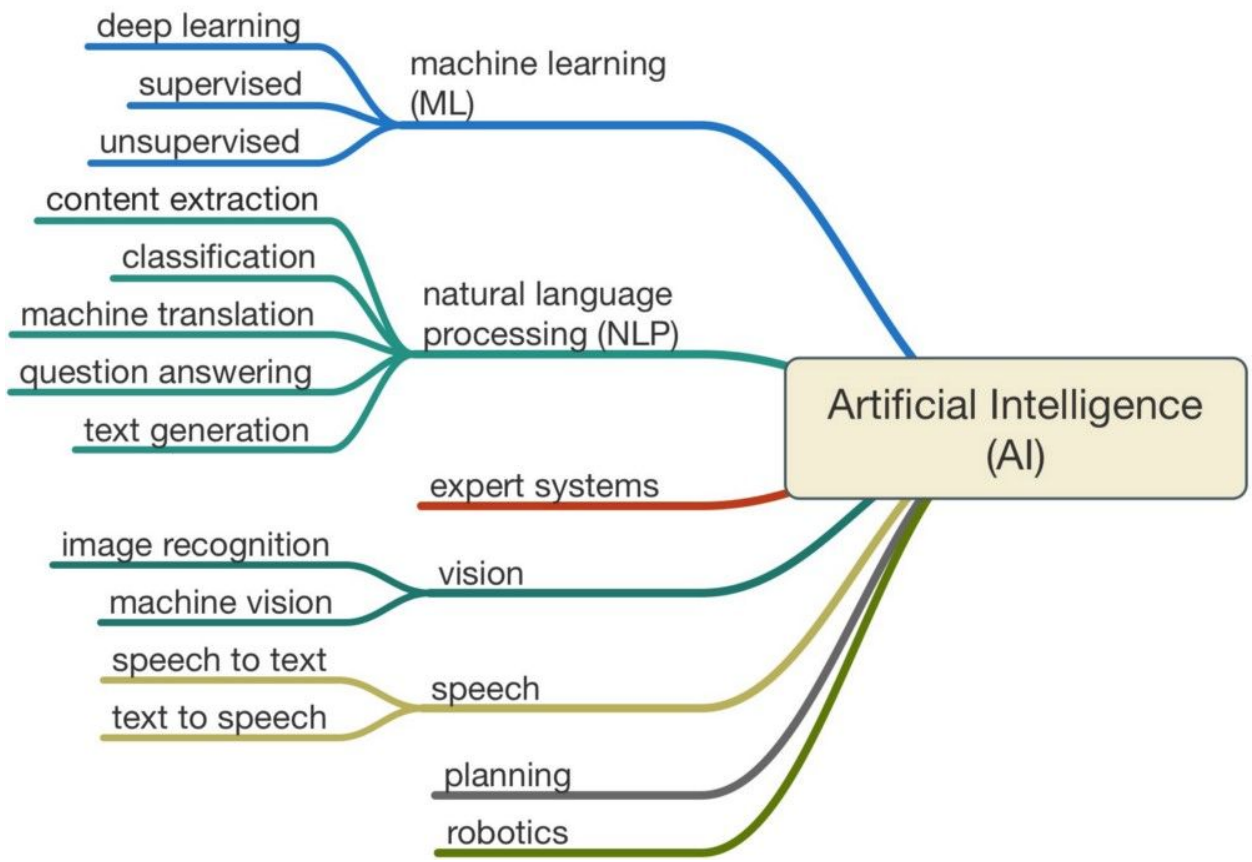


Figure 5. Hierarchy of Artificial Intelligence.
Image source: https://cdn-images-1.medium.com/max/800/1*IKS-FVwoCZkpm3MrnFLPbg.jpeg

Although ML is the focus of this paper, it should be noted that all of the AI branches can be used in Portfolio Management. For instance, Natural Language Processing (NLP) can be used to monitor social media, press releases, earnings reports and any other communication to generate signals, which may be fed into the ML training data. Vision systems are also monitoring satellite imagery to generate ML signals such as oil supply and economic activity.[13]

As shown in the figure above, ML is a subset of AI that allows systems to automatically learn and improve from experience without being explicitly programmed.[14] This contrasts with expert systems that emulate human decision making through a programed set of conditional logic, physics-based models such as computational fluid dynamics or a numerical solution to the Black-Scholes stochastic differential equation

[13]Skywatch. "4 Ways satellite imagery is changing how we invest." *Skywatch (blog)* August 22, 2016.
[14]Expert System Team. "What is Machine Learning? A definition." *Expert System (Blog)*, Accessed Mar 2019.

for options pricing. In ML, the basic architecture of the model is defined but the relationships between the features are learned automatically from the given data.

The figure below further breaks ML into sub-categories. As with the AI figure, this figure is not complete and never will be since the applications and algorithms are constantly evolving.

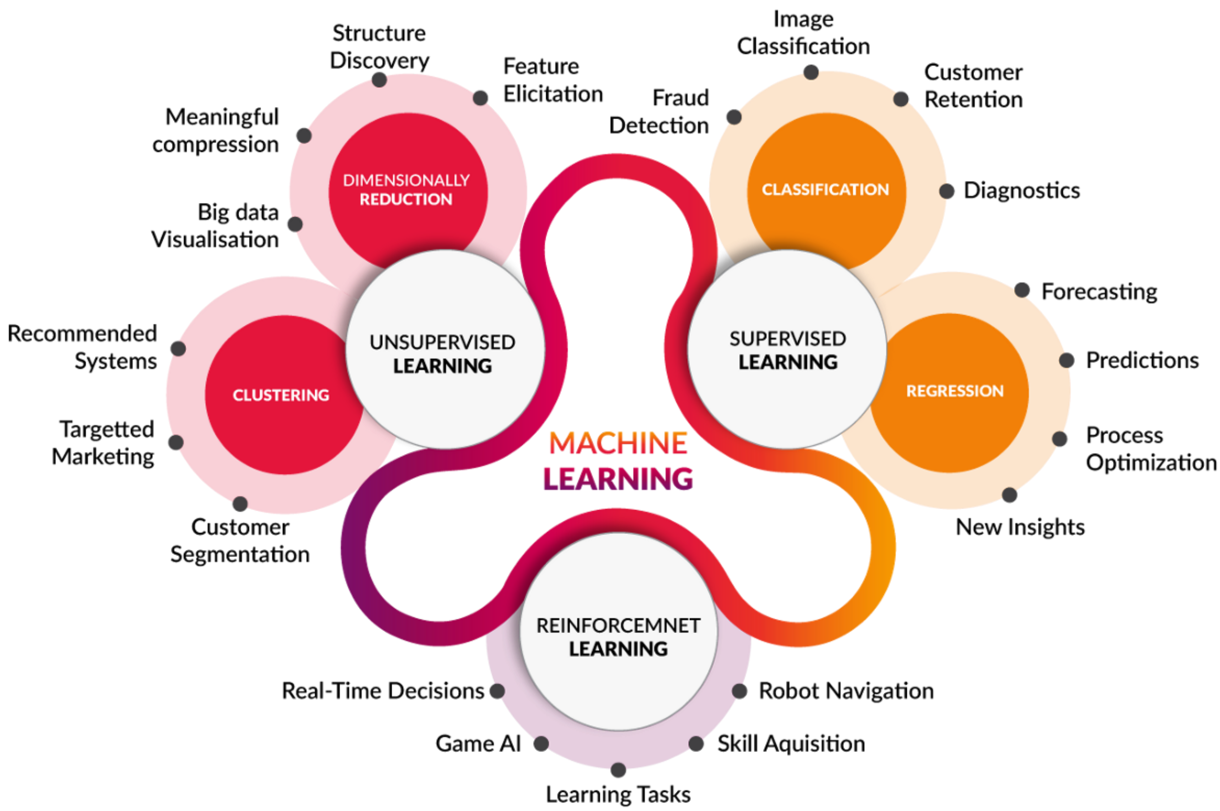


Figure 6. Hierarchy of machine Learning.
Image source: http://www.cognub.com/index.php/cognitive-platform/

Kamil Krzyk's article "Coding Deep Learning For Beginners" has an excellent high level explanation with good imagery for the three main branches of ML: Supervised, Unsupervised and Reinforcement Learning.[15]

Supervised learning, as Krzyk discusses, is probably the most recognized form of ML. It reads data with inputs and outputs identified for every point and then predicts the output of new sets of input. Note that in most ML literature the output are called labels.

---

[15] Kamil Krzyk. "Coding Deep Learning For Beginners." *Medium*, July 25 2018.

The data used in unsupervised learning doesn't include labels, and this is one of the primary distinctions between the two types of learning. Here the task is not to predict, but rather to infer the hidden structure of the unlabeled data.

Krzyk's third branch is Reinforcement Learning (RL). At a high level, RL is the process of learning the optimal strategy (***Policy***) that defines ***Actions*** that an ***Agent*** should take given the current ***State***, ***Reward*** function, and future reward ***Discount Factor***. In both supervised and unsupervised learning, the data is used to learn relationships described by the data, either between input and output (supervised) or within the input (unsupervised). RL goes a step further and tries to answer the question about what actions that the agent should have taken. For instance, if the training data had state S=1 resulting in action A=1, supervised learning would recommend A=1 for a future S=1. However, in RL it may recommend A=2 for S=1. This is based not only on the next predicted reward but also all future possible discounted rewards. It also includes the impact of that action not only on the rewards but also the future states. Think of a chess player sacrificing a queen for checkmate in three moves.

The figure below illustrates the classic RL feedback loop used to determine the optimal policy.[16] Here, the environment is modeled as a function that predicts the next reward and state given the current state and action.

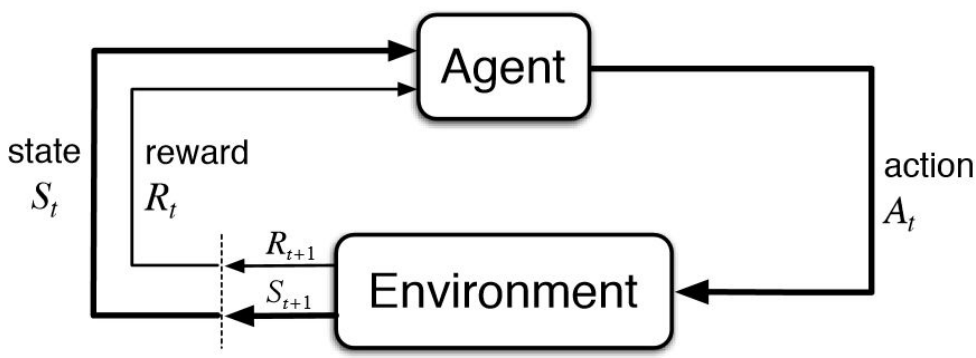


Figure 7. Reinforcement Learning feedback loop.
Image source: https://i.stack.imgur.com/eoeSq.png

As will be discussed in more depth in a later section, in a model-based RL algorithm the reward function and discount factors can be explicitly defined by the person writing the algorithm or it can be learned from the previous trade history. The actions are implicit in the algorithm definition and depend on the application. For instance the actions may be buy or sell from a defined set of assets. The state space may be hard coded to discrete values, however much of the interesting implementations use other ML routines to define these states.

[16] Steeve Huang. "Introduction to Various Reinforcement Learning Algorithms. Part I." *Medium*, Jan 12 2019.

Another sub-category missing from Figure 6 is Inverse Reinforcement Learning (IRL). Here, the reward function is not known, but we infer it from a given set of actions. For instance, if you have the trades of a successful trader you could use IRL to learn the trader's reward function, effectively what motivates them.

Note that Krzyk's article and the remainder of this paper ignores Semi-supervised learning, which uses a mix of supervised and unsupervised learning to process a combination of labeled and unlabeled data. Some ML discussions will list semi-supervised as a fourth branch, however I disagree with this. The enmeshing of these branches with ML platforms serves to capture this fourth branch, and thus does not need to be explicitly separated. For instance the state definitions in RL may use unsupervised learning to reduce thousands of signals, such as GDP, interest rates, exchange rates, earnings, dept levels and sentiment analyses of media reports into a small set of bear-to-bull market indicators.

Figures 8 and 9 come from the NYU Machine and Reinforcement Learning in Finance specialization.[17] We can see that ML can be applied to many areas in finance and the branch most applicable to portfolio management is reinforcement learning.

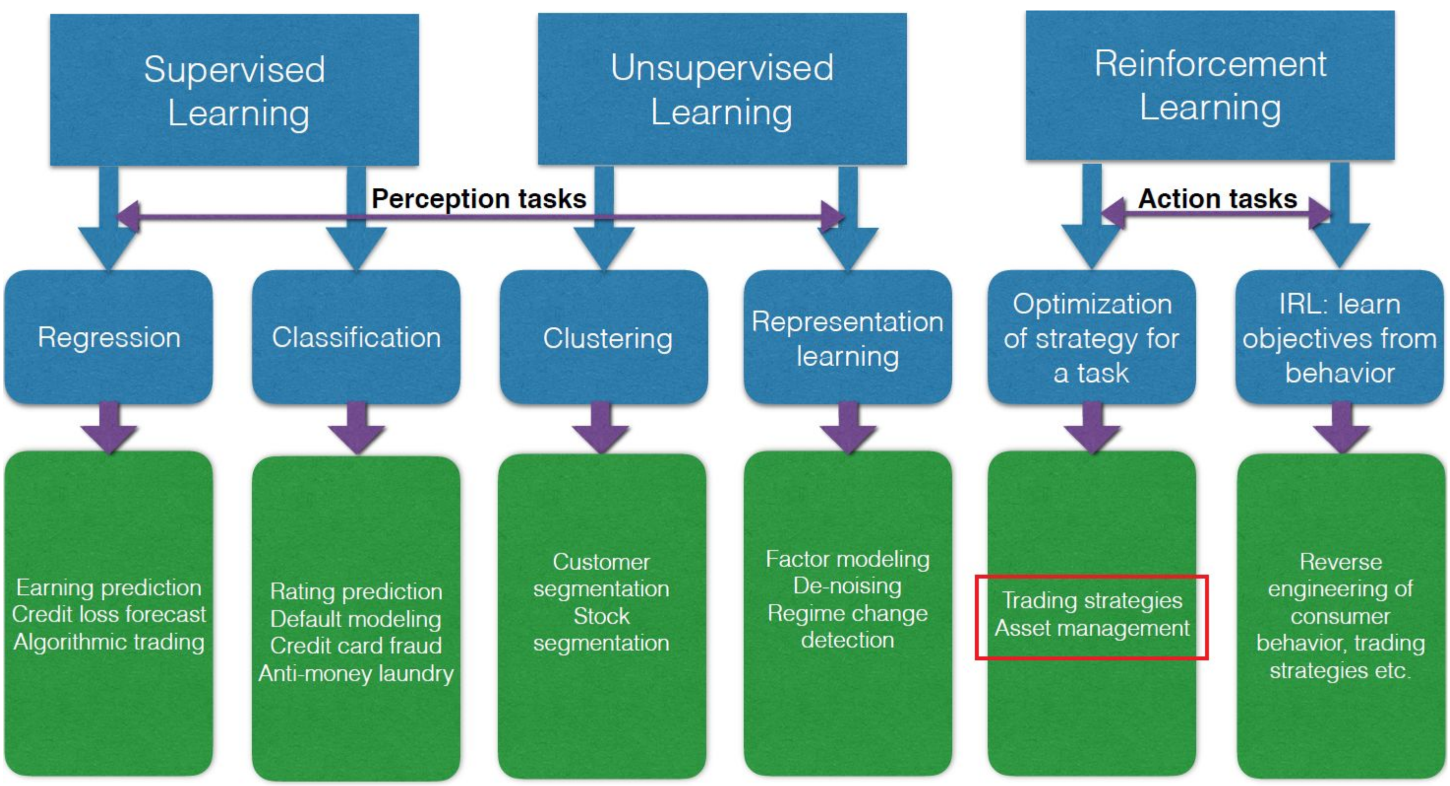


Figure 8. Machine Learning in Finance.
Image source: https://www.coursera.org/learn/guided-tour-machine-learning-finance

[17] Igor Halperin. "Guided Tour of Machine Learning in Finance." *NYU Tandon*, Coursera.

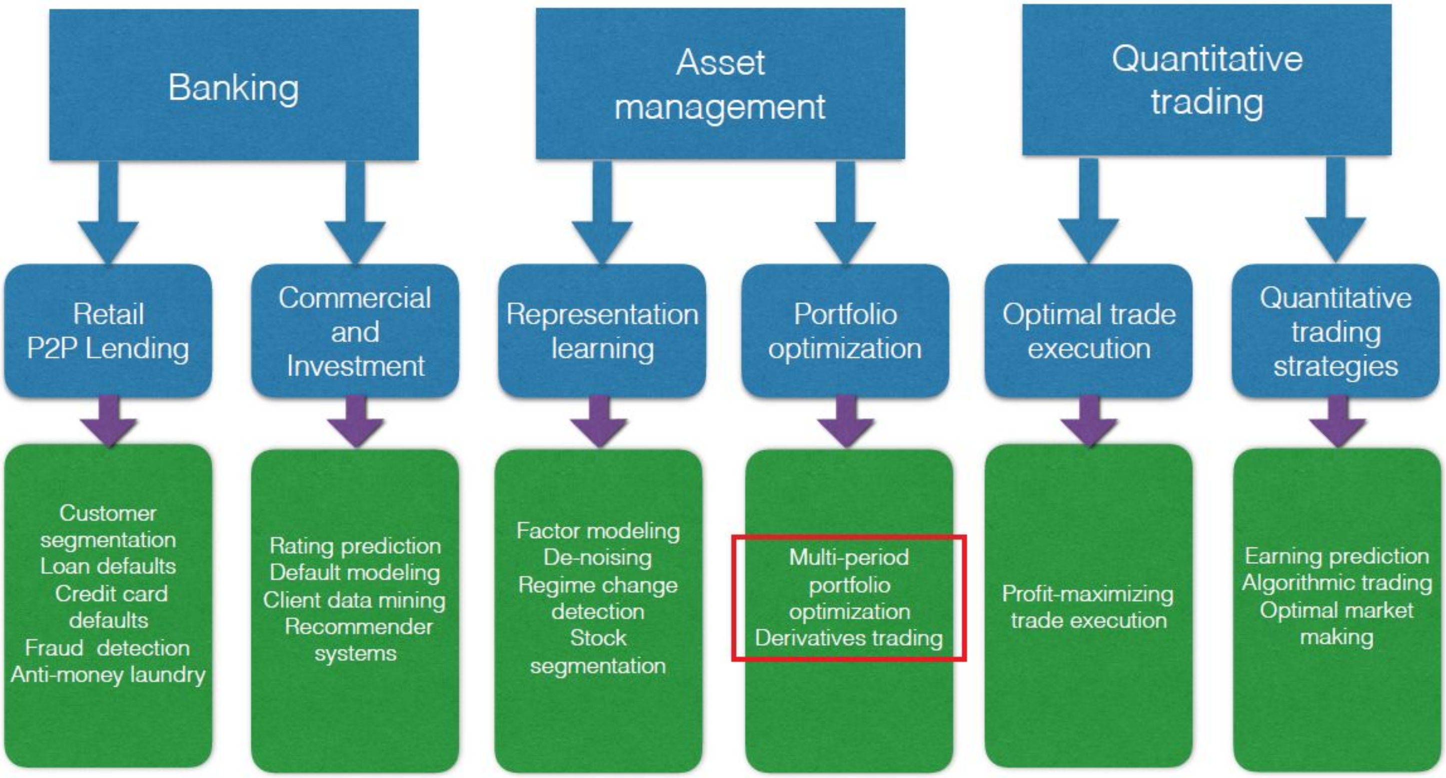


Figure 9. Machine Learning by financial application areas.
Image source: https://www.coursera.org/learn/guided-tour-machine-learning-finance

Reinforcement Learning (RL)

As discussed above, RL is one of the major branches of ML with applications to portfolio management. Note that many of the articles discussing RL are either explicitly or implicitly quoting Sutton and Barto's book which is available online. Considering Sutton and Barto's book is a 548 page introduction to RL, the treatment in this short paper is very high level. An excellent survey of books, papers, courses, conferences and blogs related to reinforcement learning was compiled by Yuri Li and is well worth reviewing.[18]

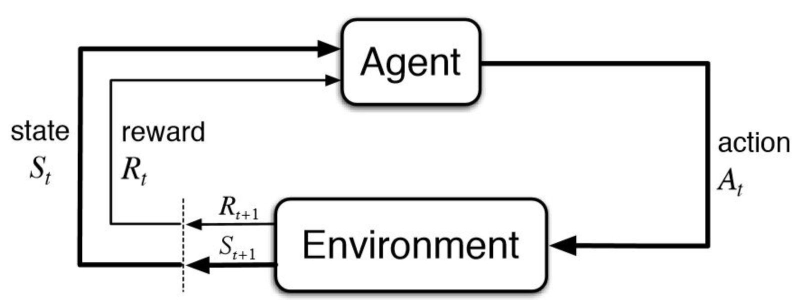
Definitions

Before we begin the discussion we need to make the following definitions.[19]

1. Agent: The algorithm or person that takes the given action.
2. Environment: The world which the agents exists. It converts the agent's current state and actions into the associated rewards and next state.
3. Action (***a***): All the possible moves that the agent can take.
4. State (***s***): Current situation returned by the environment.
5. Reward (***R***): An immediate return sent back from the environment to evaluate the last action.
6. Policy (π): The strategy that the agent employs to determine the next action based on the current state.
7. Value (***V***): The expected long-term return with discount, as opposed to the short-term reward ***R***. ***V***π(***s***) is defined as the expected long-term return of the current state ***s*** under policy π.
8. Q-value or action-value (***Q***): Similar to ***V***, except that it takes an extra parameter, the current action ***a***. ***Q***π(***s***, ***a***) refers to the long-term return of the current state ***s***, taking action ***a*** under policy π.
9. Discount Factor (γ): The discount factor applied to future rewards.

We also repeat figure 7 here to help with the definitions.

Image source: https://i.stack.imgur.com/eoeSq.png



[18] Yuxi Li. "Resources for Deep Reinforcement Learning." *Medium*, December 28 2018.

[19] Richard S. Sutton and Andrew G. Barto. Reinforcement Learning: An Introduction. 2nd edition, Cambridge, Massachusetts: *The MIT Press*, 2018.

Variants

Considering RL is a full branch of ML, it is not surprising that there are many variants and mixtures of algorithms. Here is a short list of the many RL algorithms referenced in the Bibliography material. It should be noted that not all of these algorithms are mutually exclusive. For instance, Temporal Difference (TD) learning is applied in the SARSA algorithm.

1. Q-Learning[20]
2. Fitted Q Iteration Learning[21]
3. R-Learning[22]
4. SARSA (State-Action-Reward-State-Action)
5. Modified Connectionist Q-Learning (SARSA original name)[23]
6. Posterior Sampling for Reinforcement Learning (PSRL)[24]
7. Deep Q-Networks (DQN)[25]
8. Deep Recurrent Q-Learning (DRQN)[26]
9. Deterministic Policy Gradient (DPG) Algorithms[27]
10. Deep DPG (DDPG) Algorithms[28]
11. Temporal Difference (TD) Learning[29]
12. Actor-Critic Methods[30]
13. Multi-Agent Reinforcement Learning[31]

---

[20] Christopher Watkins and Peter Dayan. "Q-Learning." *Machine Learning*, 8(3): 279–292, 1992.

[21] Damien Ernst, et al.. "Tree-based batch mode reinforcement learning." *Journal of Machine Learning Research*, 6:503–556, (2005).

[22] Anton Schwartz. "A Reinforcement Learning Method for Maximizing Undiscounted Rewards." *ICML*, 1993.

[23] G. A. Rummery and M. Niranjan. "On-Line Q-Learning using Connectionist Systems." Technical Report, *CUED/F-INFENG/TR 166*, Cambridge University, 1994.

[24] Ian Osband et al.. "(More) efficient reinforcement learning via posterior sampling." *In Proceedings of the 26th International Conference on Neural Information Processing Systems*, Volume 2 (NIPS'13), 2013.

[25] Volodymyr Mnih, et al.. "Human level control through deep reinforcement learning." *Nature*, 518(7540): 529–533, 2015.

[26] Matthew Hausknecht and Peter Stone. "Deep Recurrent Q-Learning for Partially Observable MDPs." *AAAI 2015 Fall Symposium Series, Revised*, January 11 2017.

[27] David Silver et al. "Deterministic Policy Gradient Algorithms." *31st ICML*, 2014.

[28] Timothy Lillicrap et al.. "Continuous Control With Deep Reinforcement Learning." *ICLR,* 2016.

[29] Andrew G. Barto. "Temporal difference learning." *Scholarpedia*, 2(11):1604 (2007).

[30] Pawel Wawrzynski. "Real-time reinforcement learning by sequential Actor-Critics and experience replay." *Neural networks*. 22. 1484-97, June 2019.

[31] Busoniu, L., R. Babuska, and B. De Schutter, "A comprehensive survey of multi-agent reinforcement learning." *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 38, no. 2, pp. 156–172, March 2008.

The variants above can be grouped by the following three classes. For instance, Q-Learning is an ***off-policy***, ***model-free*** algorithm than can be executed ***offline***, where as SARSA is also model-free but on-policy and most often executed online. Understanding the differences between these classes allow us to select which variants to research.

1. On-policy vs. Off-policy
2. Online vs. Offline
3. Model-based vs. Model-free

On-policy versus Off-policy[32]

All optimization algorithms face a tradeoff between ***exploration*** and ***exploitation***. If the next proposed action is always selected to maximize the value function, also known as the ***greedy*** action, it runs the risk of getting trapped in a local minimum. If it chooses to explore by selecting non-optimal actions, it may avoid the local minimum but may take much longer to converge. In the RL framework, this is expressed as off-policy and on-policy algorithms.

For off-policy algorithms, there are two distinct policies developed. The ***target policy*** is the policy being evaluated and improved to maximize the Q-value. The ***behavior policy*** is the policy driving the actions or behavior of the agent. Separating the two policies introduces freedom and simplification in the separate algorithm develop. It also makes offline methods discussed later simpler.

For on-policy algorithms, the target and behavior policies are the same. To avoid getting trapped, most on-policy algorithms will include an ***ε-greedy*** component. This results in a policy where most of the time the greedy action is selected, but there is a probability ε that a random action will be selected.

Online versus Offline Algorithms[33]

RL is inherently online. By definition, the agent interacts with the environment by taking actions and then refining the target policy based on the resulting rewards and changes in state. As mentioned in the previous section, both on-policy and off-policy algorithms must make non-optimal actions to explore as many states of the environment as possible to develop the most robust target policy. For many of the applications in the literature such as games of Go, the consequence of these exploratory actions is

[32] Richard S. Sutton and Andrew G. Barto. Reinforcement Learning: An Introduction. 2nd edition, Cambridge, Massachusetts: *The MIT Press*, 2018, p 103.

[33] Travis Mandel et al.. "Offline Evaluation of Online Reinforcement Learning Algorithms." In Proceedings of the 13th AAAI Conference on Artificial Intelligence, *AAAI Press*, February 2016, p 1926-1933.

reasonably benign. Worst case the agent loses a game and simply begins another. However, when we apply RL to manage a portfolio, these exploratory actions (trades) maybe devastating.

An offline algorithm will use historical data, such as trade history, to develop a target policy that would have performed well with the historical data. Clearly this would be required before an investor would allow the algorithm to manage their portfolio. Even after the algorithm goes online and begins learning from its own trades, it can still continue to learn in an offline mode from other traders.

Model-based versus Model-free

As discussed previously, in RL the agent takes actions based on a policy to maximize the total long-term value. This long term value is a function of both the immediate reward received from taking the next action and the sum of all expected future discounted rewards. These future rewards are a function of the next state that will result from the next action. These actions are taken in an environment that generates the next reward and forces a transition between the current and next state.

Model-based algorithms build or are given a model of this environment that can be used to plan actions and build a policy. Such a model can predict the next state (***s'***) and reward ***R***(***a***,***s***) based on the current state (***s***) and action (***a***). Note the state prediction is often probabilistic with a transition probability function ***T***(***s'***|***a***,***s***).

For instance, a classic model-based approach for option pricing uses a stochastic differential equation such as the Black-Scholes-Merton (BSM) model as the environment.[34] It then uses standard dynamic programming techniques and the Bellman optimality equation as an action-value function to identify the optimal policy.

Another model-based approach that has been shown to outperform the use of the BSM model is to use supervised machine learning such as an artificial neural network (ANN) to model the action-value function.[35]

Model-free algorithms do not build an explicit model of the environment. Instead as in the classic Q-Learning algorithm, the Q-value is estimated directly from the current state and action. There is no need to model the state transition probability function. Building an accurate transition function is very difficult and as Vladimir Vapnik, the

---

[34] Igor Halperin. "QLBS: Q-Learner in the Black-Scholes(-Merton) Worlds." *SSRN Electronic Journal*, December 2017, DOI: 10.2139/ssrn.3087076, p2.

[35] Chien-Yi Huang. "Financial Trading as a Game: A Deep Reinforcement Learning Approach." *arXiv:1807.02787,* July 8, 2018, p6.

co-inventor of the Support Vector Machine algorithm stated; "one should avoid solving more difficult intermediate problems when solving a target problem".[36]

An excellent description of the difference between model-based and model-free methods was provided by Dayan and Niv and their illustration is copied below.[37] When deciding to take the freeway or not, a Model-based method would build a model of the traffic along all possible paths, with each path being a combination of future actions. It would then calculate the total time along each path and select the path with the shortest time, which would define the next action; to take the freeway or not. After that action is taken, the model is updated with feedback from the environment and the process is repeated for the next action.

The Model-free method would simply learn that at 5:45 don't take the freeway. Here the state is the time, the action is don't take the highway and the Q-value function would be the total time to get home. The model-free approach skips the difficult task of modeling the traffic along all possible paths and learns the desired Q-value (total time) as a function of the state-action pair of 5:45 and don't take freeway.

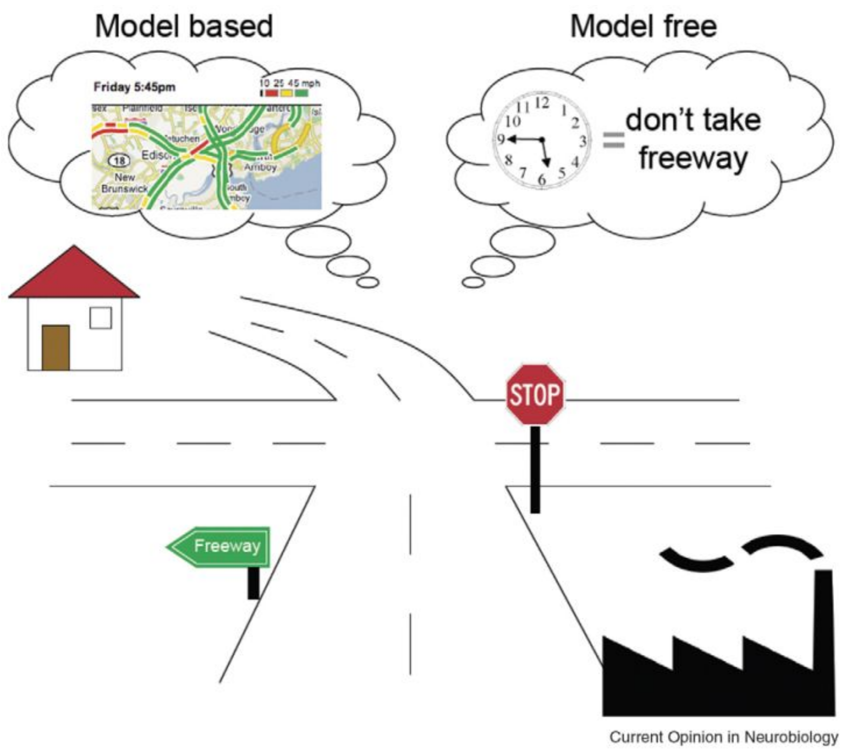


Figure 10. Model-based versus Model-free method of deciding how to get home.
Image source: https://www.princeton.edu/~yael/Publications/DayanNiv2008.pdf

[36]Vladmir N. Vapnik. "Statistical Learning Theory." *John Wiley & Sons, Inc.*, Toronto Canada, (1998).

[37]Peter Dayan and Yael Niv. "Reinforcement learning: The Good, The Bad and The Ugly." *Current opinion in neurobiology*, 18(2) 10.1016/j.conb.2008.08.003, September 2008: p 3 (Box 1).

An interesting strategy is to combine both model-based learning, which is a fast stable learner, with model-free learning, which has the capability to capture much more detail.[38]

RL for Portfolio Management

RL is a rich and deep field with an extremely active research community and an expanding list of practical applications. At its core it is an evolving system that learns to answer the question *what action should I take in a given situation based on feedback from previous actions*. The first set of celebrated RL successes came from beating world champions in games such as chess and go but has expanded into many fields such as robotics, autonomous vehicles, energy management and finance.[39] Fisher's survey of almost 50 publications related to RL in financial markets emphases the many different applications within finance alone.[40] If you can define a state space, a set of possible actions and the rewards, you can build a RL algorithm with varying degrees of success. The three common applications in finance are portfolio optimization, market making and optimal trade execution.[41]

This section will focus on portfolio management with a state space defined on the following page. The proposed actions are buy, sell or hold from defined set of *n* assets at a given frequency. The reward is defined as a increase in portfolio value with a given level of risk and the total Q-value to be maximized is the sum of expected discounted rewards out to a given time horizon.

---

[38] Anusha Nagabandi et al.. "Neural Network Dynamics for Model-Based Deep Reinforcement Learning with Model-Free Fine-Tuning." *2018 IEEE ICRA*, DOI: 10.1109/ICRA.2018.8463189, May 25 2018.

[39] Li, Yuxi. "Reinforcement Learning Applications." *Medium*, October 15 2018.

[40] Thomas G. Fischer "Reinforcement learning in financial markets - a survey." *FAU Discussion Papers in Economics*, ISSN 1867-6707, 2018.

[41] Igor Halperin. "What Are The Latest Works On Reinforcement Learning In The Financial Field?" *Forbes*, July 25 2018.

State Definition

Like manual portfolio management, the state definition greatly influences the success of the policy. A possible state space could contain the following, where *n* is the number of assets in the possible investment universe.[42]

Position (n): The value of the assets in the portfolio.

Correlations ((n²+n)/2): This is the correlation between the possible assets including the standard deviation of each asset.

Time features (3): Time of day, day of week and day of year to capture temporal trends.

Technical indicators (7*n): Parabolic Stop And Reverse (SAR), Moving Average Convergence Divergence (MACD), Relative Strength Index (RSI), Ichimoku Kinko Hyo, Bollinger Bands, stochastic indicator and Average Directional Index (ADX).[43]

Fundamental Indicators (5*n): Five-Factor Fama-French model.[44]

Economic Indicators (n + 6): Sector, seasonally adjusted GDP growth, Consumer Price Index (CPI), Industrial Production Index (IPI), interest rate and unemployment rate.[45]

Note that the number of dimensions in the above state space is $8+15.5n+\frac{1}{2}n^2$. If we were considering 100 possible stocks, the space would have 6,558 dimensions! However, 5,050 come from the correlations. Since these correlations may be captured in the technical and fundamental indicators, we could eliminate them to get 1,508 dimensions.

Clearly choosing the most effective state definition is critical and requires a strong understanding of the markets. Using RL may provide a trading advantage over traditional methods, however the efficient market hypothesis, which states that prices reflect all available information making superior returns impossible, suggests that incorporating unique state dimensions will always be vital. Effectively building state

---

[42] Chien-Yi Huang. "Financial Trading as a Game: A Deep Reinforcement Learning Approach." *arXiv:1807.02787,* July 8, 2018, p. 4-5.

[43] Harry Nicholls. "7 Popular Technical Indicators and How to Use Them to Increase Your Trading Profits." *The Medium*, April 8 2018.

[44] Eugene F. Fama and Kenneth R. French. "A five-factor asset pricing model." *Journal of Financial Economics*, Vol 116, Issue 1, p1-22, April 2015.

[45] Francisco Jareño and Loredana Negrut. "US Stock Market and Macroeconomic Factors." *Journal of Applied Business Research*, Vol 32. p325-340. 10.19030/jabr.v32i1.9541, February 2016.

dimensions with advanced ML, incorporates information not available to other investors, thereby avoiding the limits of the efficient market hypothesis.

As mentioned, other ML algorithms may be used to define the state. For instance, unsupervised clustering and dimensional reduction algorithms may process a high dimensional space to provide a much smaller and more manageable state space. Natural language processing can monitor media and earnings reports to generate a sentiment state dimension and image processing can monitor satellite imagery to gauge economic activity or production output. In general, as the amount of and access to data increases, ML can generate signals that can be incorporated into an evolving reinforcement algorithm.

On-policy versus Off-policy

Reviewing the distinction between on-policy and off-policy algorithms, it becomes clear that off-policy algorithms naturally fit with the offline requirements of a ML algorithm applied to portfolio management. The behaviour policy could be data from both the RL algorithm and other traders, and the target policy development could be performed rapidly in batch mode.

Model-based versus Model-free

The literature is very active with both model-based and model-free methods and it isn't clear which is more attractive for portfolio management. However, methods based on Q-Learning that are both model-free and off-policy have been shown to perform very well for option pricing as well as financial trading systems.[46,47] SARSA is another very popular model-free algorithm, however it is a on-policy routine and on-policy may not be as effective for portfolio management.

---

[46] Igor Halperin. "The QLBS Q-Learner Goes NuQLear: Fitted Q Iteration, Inverse RL, and Option Portfolios." *SSRN Electronic Journal*, January 2018.

[47] Marco Corazza and Francesco Bertoluzzo. "Q-Learning-Based Financial Trading Systems with Applications". *University Ca' Foscari of Venice*, Dept. of Economics Research Paper Series No. 15 ISSN: 1827-3580, Oct 2014.

Existing Machine Learning Applications

A quick review of the bibliography illustrates that the research into ML and portfolio management is very active, however it is difficult to determine its penetration into the finance industry. This isn't surprising considering the competitive nature of the investment industry and the recent emergence of practical reinforcement learning. Having an effective RL trading system would be a major competitive advantage that investors would not want to share. Once it becomes more mainstream, we may see more industrial publications.

The following organizations have publicised ML platforms, however it is difficult to determine their exact methods.

Taaffeite Capital Management [48]

The documentation available for Taaffeite Capital Management is limited, however their website indicates that they are using machine learning as a pattern recognition tool to identify mispriced assets and optimize trading.

TWO SIGMA

Two Sigma is another company that appears to be utilizing machine learning based on their website but the information is very limited. They state that they use machine learning to find connections in the world's data.[49]

Goldman Sachs: 5 Next Wave of Innovation ETFs [50] [51]

Just this March, Goldman Sachs and Motif Inc. launched five new ETFs "built with machine learning". The websites and Wall Street Journal article don't give specific detailed of the algorithms but they do mention that Motif processes patent databases, academic journals and company financial reports to help build the ETFs. This suggests NLP and traditional data mining to classify themed ETFs such as the Data-Driven World (GDAT) ETF.

---

[48] Taaffeite. "Taaffeite Capital Management." *taaffeitecm.com*.

[49] Two Sigma. "Approach." *twosigma.com*.

[50] Asjylyn Loder. "Goldman Rolls Out New ETFs Focused On Artificial Intelligence." *The Wall Street Journal*, March 2019.

[51] Motif. "The next wave of innovation is here." *motif.com*.

Bank of America Merrill Lynch [52] [53]

The Bank of America Merrill Lynch states that their "MosaicAlgo™ is a machine learning algorithm that utilizes order and stock characteristics together with TCA data to determine the optimal strategy for execution." Note that TCA refers to a transaction cost analysis. Trade execution is an active topic in the machine learning literature.[54] Although they don't indicate what algorithms are used in the MosaicAlgo, developing an optimal strategy for execution aligns very well with reinforcement learning, which develops an optimal policy (strategy).

EquBot [55]

EquBot uses the IBM Watson computing platform to develop an AI ETF. Similar to Motif, they are scanning "millions of articles and news sources" to gather information. This is very similar to the state definition discussion in a previous section. It is interesting that they are modeling over 15,000 global companies. This makes for an extremely highly dimensional state space. They don't indicate how they process this data to make it manageable. It could be as simple as a basic filtering to down sample or a more sophisticated unsupervised method or most likely a combination of methods. They are also combining both fundamental and quantitative analysis as discussed in the state space section.

J.P. Morgan

J.P. Morgan has released more quality information about their ML algorithms than anyone else and has a very mature system. For instance, Bank of America Merrill Lynch states they have a ML algorithm for trade execution but doesn't provide details, however J.P. Morgan is explicitly stating that they have a RL algorithm. In particular they have a Certainty Equivalent Reinforcement Learning algorithm that incorporates individual investor risk tolerance into the discount factor applied to future rewards.[56]

J.P. Morgan is also using NLP and unsupervised clustering algorithms to generate state space dimensions. In one study they applied NLP to 250,000 analyst reports and 100,000 news articles to assist in equity investment decisions.[57] In a 2017 report, they

---

[52] Bank of America Merrill Lynch. "Instinct® Equities Electronic Trading Guide." *campus.bankofamerica.com*.

[53] Bank of America Merrill Lynch. "Optimizing Execution for Asian Equities." *bofaml.com*.

[54] Dixon, Matthew. "Sequence classification of the limit order book using recurrent neural networks." *Journal of Computational Science*, Vol. 24, January 2018: 227-286.

[55] EquBot. "Our Technology." equbot.com.

[56] Sarah Butcher. "JPMorgan's new guide to machine learning in algorithmic trading." *eFinancialCarrers*, December 3 2018.

[57] J.P. Morgan. "Innovations in Finance with Machine Learning, Big Data and Artificial Intelligence." *jpmorgan.com*.

include a full section called "Handbook of Alternative Data" and the appendix even includes techniques for data collection from websites.[58]

In a paper published by J.P. Morgan, they describe a very novel Hierarchical RL approach that consists of layers of policies with different decision frequency.[59] They also make an interesting comparison between a game of Go, which has about 200 steps per game, and medium frequency electronic trading, which has about 3600 steps per hour. This illustrates the complexity of the trading problem. Although they didn't described the inner workings of their algorithm, they did mention that they bypassed the problem of how the world works and built policies that selected the best action directly from the data. This is clearly describing a model-free approach.

In another study of the interest rate market, they trained the routine with 1,250 features from 2000 to 2016 (offline behaviour policy) and applied it to trading in 2017, with the results shown below.[60] It clearly shows the RL methods outperforming the standard trading methods (systematic). It also shows a trend that has been highlighted in other studies, where ML is extremely effective in intraday trading, moderately effective in medium term investing and not as effective in long-term investing where unexpected events can drive the market.

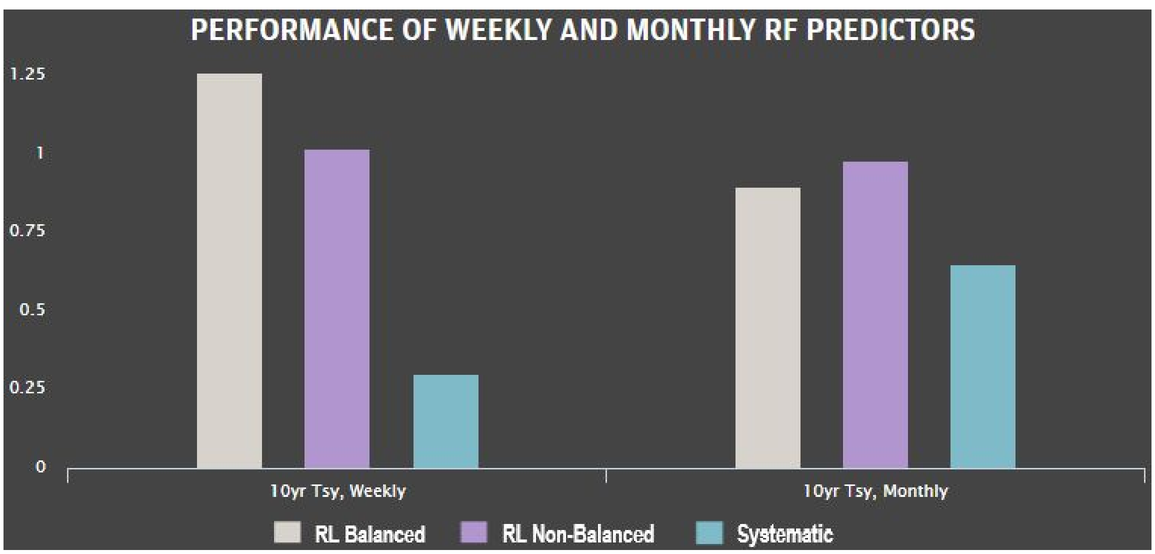


Figure 11. J.P. Morgan November 22 2017 RL performance report.
Image source: https://www.jpmorgan.com/global/research/machine-learning

---

[58] Kolanovic, Marko and Rajesh T. Krishnamach. "Big Data and AI Strategies: Machine Learning and Alternative Data Approach to Investing." *J.P. Morgan Global Quantitative & Derivatives Strategy,* May 18 2017.

[59] Vangelis Bacoyannis, Vacslav Glukhov, Tom Jin, Jonathan Kochems and Doo Re Song. "Idiosyncrasies and challenges of data driven learning in electronic trading." *arXiv:1811.09549v2*, November 30 2018.

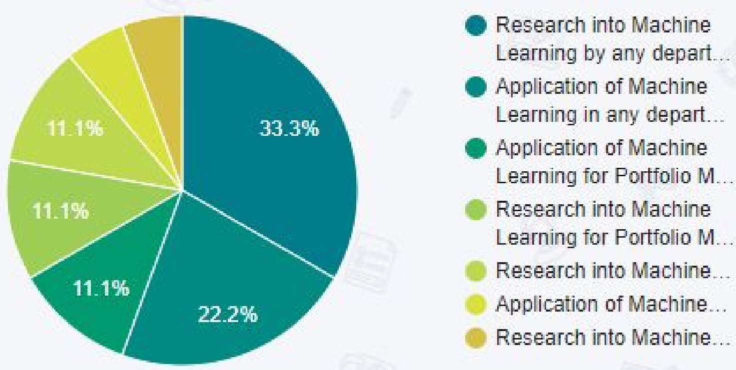
[60] J.P. Morgan. "Innovations in Finance".

Survey Results

As part of this research a small survey was created to gather information about ML applications in Portfolio Management in the Toronto, Canada area. Unfortunately the number of responses to this survey was limited, with only 11 responses. Keeping this small sample size in mind, the results are shown below.

1. Please check the following items that you are aware of within your company?

(9 responses)

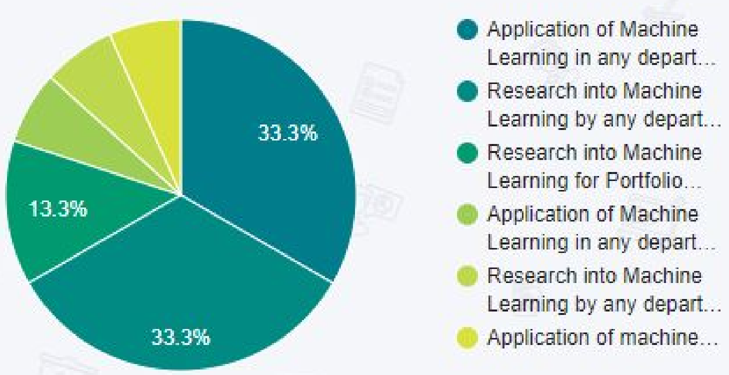


Summary

Answer	Percentage	Responses
Research into Machine Learning by any department	33.3%	6
Application of Machine Learning in any department	22.2%	4
Application of Machine Learning for Portfolio Management	11.1%	2
Research into Machine Learning for Portfolio Management	11.1%	2
Research into Machine Learning for Portfolio Management	11.1%	2
Application of Machine Learning in any department	5.6%	1
Research into Machine Learning by any department	5.6%	1

2. Please check the following items that you are aware of within other financial institutions?

(9 responses)

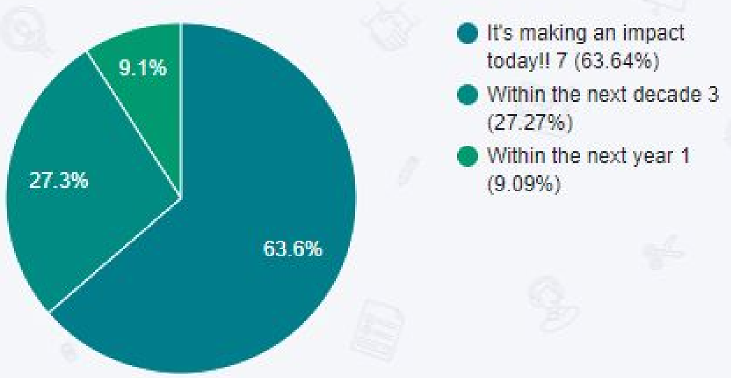


Summary

Answer	Percentage	Responses
Application of Machine Learning in any department	33.3%	5
Research into Machine Learning by any department	33.3%	5
Research into Machine Learning for Portfolio Management	13.3%	2
Application of Machine Learning in any department	6.7%	1
Research into Machine Learning by any department	6.7%	1
Application of machine learning for Portfolio Management	6.7%	1

3. When do you think Machine Learning will impact Portfolio Management ?

(11 responses)

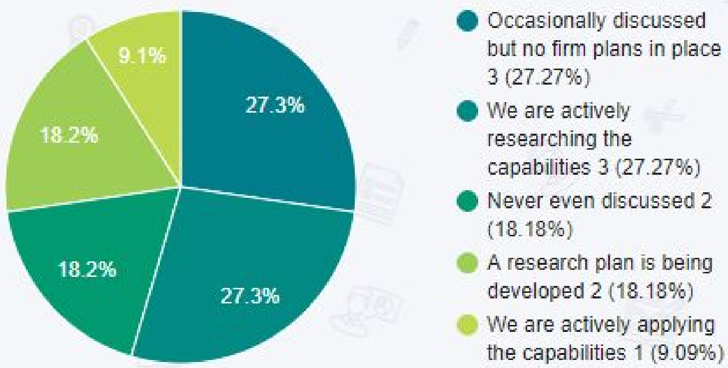


Summary

Answer	Percentage	Responses
It's making an impact today!!	63.64%	7
Within the next decade	27.27%	3
Within the next year	9.09%	1

4. What is your organization's level of interest in Machine Learning for Portfolio Management?

(11 responses)

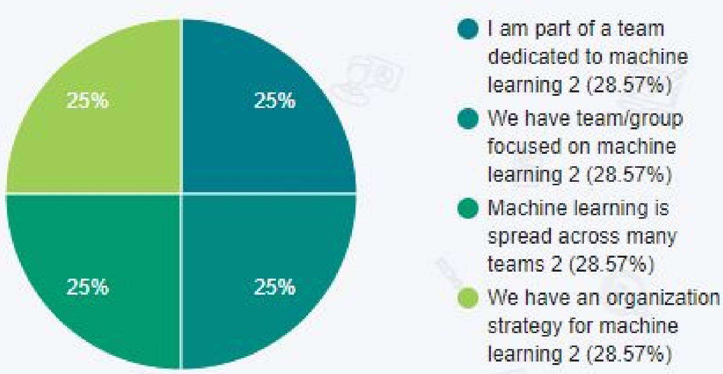


Summary

Answer	Percentage	Responses
Occasionally discussed but no firm plans in place	27.27%	3
We are actively researching the capabilities	27.27%	3
Never even discussed	18.18%	2
A research plan is being developed	18.18%	2
We are actively applying the capabilities	9.09%	1

5. Please check the following with respect to Portfolio Management within your organization?

(7 responses)



Summary

Answer	Percentage	Responses
I am part of a team dedicated to machine learning	25%	2
We have team/group focused on machine learning	25%	2
Machine learning is spread across many teams	25%	2
We have an organization strategy for machine learning	25%	2

For the free form question, "Please describe Machine Learning applications or research within your organization", we got the following three interesting responses.

> *"Models for Portfolio Strategy"*
>
> *"We are applying machine learning to improve portfolio selection and portfolio risk management over traditional methods."*
>
> *"We are researching many ways ML can be applied within our organization with pockets of applications."*

Even from this limited sample size we can see that ML is considered to hold potential to make an impact in portfolio management, but it is still in its infancy. An optimist would view this as an opportunity to gain an advantage over the competition before the techniques become more common.

To supplement this survey results are shown below from a similar survey published by J.P. Morgan in 2017.[61] This survey shows over 20% adoption but unfortunately they use the term "Artificial Intelligence", which could be applied to an ecosystem of methods.

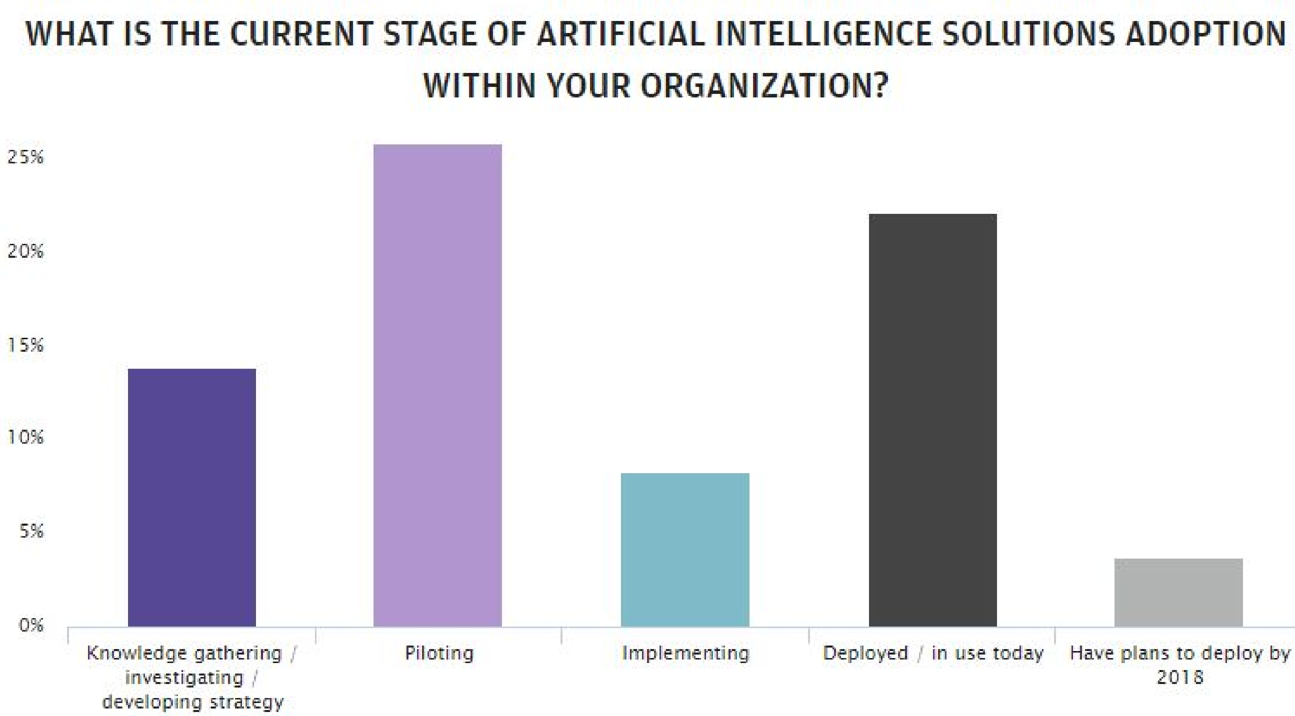


Figure 12. J.P. Morgan survey of AI adoption.
Image source: https://www.jpmorgan.com/global/research/machine-learning

---

[61] J.P. Morgan. "Innovations in Finance with Machine Learning, Big Data and Artificial Intelligence." *jpmorgan.com*.

Conclusions and Future Work

The motivation of this research was to gain a deeper understanding of how machine learning can be applied to financial portfolio management. Through an exploration of existing literature, primary research, and analysis of the market as it currently exists, one observation is clear: Portfolio Management will benefit from ML.

Supervised and unsupervised learning can be applied tactically to generate signals and define the state of the investment environment. This information can then be used within a Reinforcement Learning (RL) framework to develop a policy that determines what actions (trades) should be taken. The research community is extremely active developing new RL algorithms and possibly more importantly combining existing techniques into novel ways to create an integrated system. The best example from the investment industry was J.P. Morgan's Hierarchical RL platform, which can execute on spectrum from high frequency trading to time horizons of months.

The next step in this research will be to build a RL platform to begin experimenting with different signals and algorithms. It will be an off-policy and model-free method such as Deep Q-Learning (DQN). DQN is a good initial algorithm due to the extensive related literature and tutorials. This ensures a stable and effective platform can be generated quickly and provides a baseline for comparison. An online program such as Udacity's Nanodegree program "Deep Reinforcement Learning" will also be used to accelerate the learning of this approach.[62]

With an initial system in place, the avenues of research and application quickly multiply. This includes examining which signals should be added to the state space and what ML algorithms can generate these signals. The comparison of different RL algorithms is also a large and extremely active area of research. There are also a large number of application areas to test such as option pricing, order-book management, high-frequency trading, ETF management and general portfolio management. An operational platform also provides an excellent education resource and source of inspiration for other students.

RL is an extremely exciting area of research and holds the promise of revolutionizing the financial industry. To stay competitive we must continue the research into RL and build a RL platform for further research and development.

[62]Udacity. "Deep Reinforcement Learning." Nanodegree program. https://www.udacity.com/course/deep-reinforcement-learning-nanodegree--nd893

Bibliography

Ashraf, Mohammad. "Reinforcement Learning Demystified: A Gentle Introduction." *Towards Data Science*. (April 7 2018). https://towardsdatascience.com/reinforcement-learning-demystified-36c39c11ec14

Ashraf, Mohammad. "Reinforcement Learning Demystified: Solving MDPs with Dynamic Programming." *Towards Data Science*. (May 18 2018). https://towardsdatascience.com/reinforcement-learning-demystified-solving-mdps-with-dynamic-programming-b52c8093c919

Bacoyannis, Vangelis, Vacslav Glukhov, Tom Jin, Jonathan Kochems, Doo Re Song. "Idiosyncrasies and challenges of data driven learning in electronic trading." *arXiv:1811.09549v2*, November 30 2018. https://arxiv.org/pdf/1811.09549v2.pdf

Bank of America Merrill Lynch. "Instinct® Equities Electronic Trading Guide." Accessed March 21 2019. https://campus.bankofamerica.com/content/dam/boamlimages/documents/PDFs/electronic_trading_guide.pdf

Bank of America Merrill Lynch. "Optimizing Execution for Asian Equities." Accessed March 21 2019. https://www.bofaml.com/en-us/content/algorithmic-trading-strategies-asian-stock-market.html

Barto, Andrew G. "Temporal difference learning." *Scholarpedia*, 2(11):1604 (2007). http://www.scholarpedia.org/article/Temporal_difference_learning

Beceiro, Rodrigo. "What is Artificial Intelligence and why now?" *Medium* (August 8 2018). https://becominghuman.ai/what-is-ai-and-why-now-79d94f77dc91

Bertsimas, Dimitris, Vishal Gupta and Ioannis Ch. Paschalidis. "Inverse Optimization: A New Perspective on the Black-Litterman Model." *Operations Research* Vol. 60, No. 6 (November 2012): 1389-1403. https://doi.org/10.1287/opre.1120.1115

Black, Fischer and Myron Scholes. "The Pricing of Options and Corporate Liabilities." *Journal of Political Economy*. Vol. 81, issue 3 (1973): 637–654. doi:10.1086/260062.

Black, Fischer and Robert Litterman. "Global Portfolio Optimization." *Financial Analysts Journal* Vol. 48, Issue 3 (October 1992): 28-43. https://www.cfapubs.org/doi/abs/10.2469/faj.v48.n5.28

Bloch, Daniel Alexandre. "Machine Learning: Models And Algorithms." *SSRN Electronic Journal* (February 2019). Available at SSRN: https://ssrn.com/abstract=3307566

Boyd, Stephen, Enzo Busseti, Steven Diamond, Ronald N. Kahn, Kwangmoo Koh, Peter Nystrup and Jan Speth. "Multi-Period Trading via Convex Optimization." *Foundations and Trends® in Optimization* (© 2017 now Publishers Inc.). DOI: 10.1561. https://web.stanford.edu/~boyd/papers/pdf/cvx_portfolio.pdf

Bughun, Jacques, Eric Hazan, Sree Ramaswamy, Michael Chui, Tera Allas, Peter Dahlström, Nicolaus Henke and Monica Trench. "Artificial Intelligence The Next Digital Frontier?" *McKinsey Global Institute*, Discussion Paper (June 2017). https://www.mckinsey.com/~/media/McKinsey/Industries/Advanced%20Electronics/Our%20Insights/How%20artificial%20intelligence%20can%20deliver%20real%20value%20to%20companies/MGI-Artificial-Intelligence-Discussion-paper.ashx

Busoniu, L., R. Babuska, and B. De Schutter, "A comprehensive survey of multi-agent reinforcement learning." *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, vol. 38, no. 2, pp. 156–172, Mar. 2008. http://www.dcsc.tudelft.nl/~bdeschutter/pub/rep/07_019.pdf

Butcher, Sarah. "JPMorgan's new guide to machine learning in algorithmic trading." *eFinancialCarrers*, (December 3 2018). https://news.efinancialcareers.com/ca-en/329751/jpmorgans-new-guide-to-machine-learning-in-algorithmic-trading

Campbell, Rob. "Defensiveness: A suitcase word." *Mawer*, (July 18 2018). https://www.mawer.com/the-art-of-boring/blog/defensiveness-a-suitcase-word/

Columbus, Louis. "Roundup Of Cloud Computing Forecasts, 2017." *Forbes*, (April 2017). https://www.forbes.com/sites/louiscolumbus/2017/04/29/roundup-of-cloud-computing-forecasts-2017/#7799e1fd31e8

Corazza, Marco and Andrea Sangalli. "Q-Learning and SARSA: A Comparison between Two Intelligent Stochastic Control Approaches for Financial Trading." *University Ca' Foscari of Venice, Dept. of Economics Research Paper Series No. 15* ISSN: 1827-3580 (2015). https://ssrn.com/abstract=2617630

Corazza, Marco and Francesco Bertoluzzo. "Q-Learning-Based Financial Trading Systems with Applications". *University Ca' Foscari of Venice, Dept. of Economics Research Paper Series No. 15* ISSN: 1827-3580 (October 2014). https://ssrn.com/abstract=2507826

Corazza, Marco and Francesco Bertoluzzo. "Reinforcement Learning for automatic financial trading: Introduction and some applications". *University Ca' Foscari of Venice, Dept. of*

Economics Research Paper Series No. 33 ISSN: 1827-3580 (December 2012). https://ssrn.com/abstract=2192034

Cont, Rama. "Statistical Modeling of High Frequency Financial Data: Facts, Models and Challenges." *University of Oxford-CNRS* (March 1, 2011). https://ssrn.com/abstract=1748022.

Dayan, Peter and Yael Niv. "Reinforcement learning: The Good, The Bad and The Ugly." *Current opinion in neurobiology*. 18(2): 185-96. 10.1016/j.conb.2008.08.003 (September 2008). https://www.princeton.edu/~yael/Publications/DayanNiv2008.pdf

Dixon, Matthew. "Sequence classification of the limit order book using recurrent neural networks." *Journal of Computational Science*, Vol. 24 (January 2018): 227-286. https://doi.org/10.1016/j.jocs.2017.08.018

EquBot. "Our Technology." Accessed march 21 2019. https://equbot.com/technology/

Ernst, Damien, Louis Wehenkel and Pierre Geurts. "Tree-based batch mode reinforcement learning." *Journal of Machine Learning Research*, 6:503–556, (2005). http://www.jmlr.org/papers/volume6/ernst05a/ernst05a.pdf

Expert System Team. "What is Machine Learning? A definition." *Expert System (Blog)*, Accessed March 3 2019. https://www.expertsystem.com/machine-learning-definition/

Fama, Eugene F. and Kenneth R. French. "A five-factor asset pricing model." *Journal of Financial Economics*, Vol 116, Issue 1, p1-22, (April 2015). https://doi.org/10.1016/j.jfineco.2014.10.010

Fellah, David. "Active Learning in Trading Algorithms." *EMEA Linear Quant Research Group at J.P. Morgan* (Slides for QuantCon Singapore November 2016). https://www.quantopian.com/posts/quantcon-singapore-2016-presentations

Fischer, Thomas G. "Reinforcement learning in financial markets - a survey." *FAU Discussion Papers in Economics*, ISSN 1867-6707 (2018). http://hdl.handle.net/10419/183139

Fumo, David. "Types of Machine Learning Algorithms You Should Know." *Towards Data Science*, (June 2017). https://towardsdatascience.com/types-of-machine-learning-algorithms-you-should-know-953a08248861

Gunst, Carole. "10 Eye-opening Stats About the Growth of Big Data." *Attunity*, (August 2018). https://www.attunity.com/blog/10-eye-opening-stats-about-the-growth-of-big-data/

Halperin, Igor. “Guided Tour of Machine Learning in Finance.” *New York University Tandon School of Engineering*, (Online Course & Material). https://www.coursera.org/learn/guided-tour-machine-learning-finance

Halperin, Igor. “Fundamentals of Machine Learning in Finance.” *New York University Tandon School of Engineering*, (Online Course & Material). https://www.coursera.org/learn/fundamentals-machine-learning-in-finance

Halperin, Igor. “Reinforcement Learning in Finance.” *New York University Tandon School of Engineering*, (Online Course & Material). https://www.coursera.org/learn/reinforcement-learning-in-finance

Halperin, Igor. “Overview of Advanced Methods of Reinforcement Learning in Finance.” *New York University Tandon School of Engineering*, (Online Course & Material). https://www.coursera.org/learn/advanced-methods-reinforcement-learning-finance

Halperin, Igor. “The QLBS Q-Learner Goes NuQLear: Fitted Q Iteration, Inverse RL, and Option Portfolios.” *SSRN Electronic Journal* (January 2018). https://arxiv.org/abs/1801.06077

Halperin, Igor. “QLBS: Q-Learner in the Black-Scholes(-Merton) Worlds.” *SSRN Electronic Journal*, (December 2017) DOI: 10.2139/ssrn.3087076. https://arxiv.org/abs/1712.04609

Halperin, Igor. “What Are The Latest Works On Reinforcement Learning In The Financial Field?” *Forbes* (July 25 2018). https://www.forbes.com/sites/quora/2018/07/25/what-are-the-latest-works-on-reinforcement-learning-in-the-financial-field/#40f8e12344df

Halperin, Igor and Ilya Feldshteyn. “Market Self-Learning of Signals, Impact and Optimal Trading: Invisible Hand Inference with Free Energy.” *eprint arXiv:1805.06126* (May 2018). https://arxiv.org/abs/1805.06126

Halperin, Igor and Matthew Dixon. “"Quantum Equilibrium-Disequilibrium": Asset Price Dynamics, Symmetry Breaking, and Defaults as Dissipative Instantons.” (August 2018). https://arxiv.org/abs/1808.03607

Han, James, Johnny Hong, Nicholas Sutardja and Sio Fong Wong. “Machine learning techniques for price change forecast using the limit order book data.” (December 2015). http://jcyhong.github.io/assets/machine-learning-price-movements.pdf

Hausknecht, Matthew and Peter Stone. “Deep Recurrent Q-Learning for Partially Observable MDPs.” Association for the Advancement of Artificial Intelligence (AAAI), 2015 Fall Symposium Series, Revised (January 11 2017). https://arxiv.org/pdf/1507.06527v4.pdf

Hens, Thorsten, and Peter Woehrmann. “Strategic Asset Allocation and Market Timing: A Reinforcement Learning Approach.” *eprint arXiv:1808.03607* (January 27, 2006). https://ssrn.com/abstract=883595

Hodjat, Babak. “The AI Resurgence: Why Now?” *Wired*, (March 2015). https://www.wired.com/insights/2015/03/ai-resurgence-now/

Huang, Chien-Yi. “Financial Trading as a Game: A Deep Reinforcement Learning Approach.” arXiv:1807.02787 (July 8, 2018). https://arxiv.org/pdf/1807.02787.pdf

Huang, Steeve. “Introduction to Various Reinforcement Learning Algorithms. Part I (Q-Learning, SARSA, DQN, DDPG).” *Medium*, (January 12 2019). https://towardsdatascience.com/introduction-to-various-reinforcement-learning-algorithms-i-q-learning-sarsa-dqn-ddpg-72a5e0cb6287

J.P. Morgan. “Innovations in Finance with Machine Learning, Big Data and Artificial Intelligence.” jpmorgan.com, Accessed March 22 2019. https://www.jpmorgan.com/global/research/machine-learning

Jareño, Francisco and Loredana Negrut. “US Stock Market and Macroeconomic Factors.” *Journal of Applied Business Research*, Vol 32. p325-340. 10.19030/jabr.v32i1.9541 (February 2016). https://previa.uclm.es/profesoradO/fjareno/DOCS/9541-36008-2-PB.pdf

Kaul, Aditya and Clint Wheelock. “Artificial Intelligence Market Forecasts.” *Tractica Research Report*. (3Q 2016). https://www.tractica.com/wp-content/uploads/2016/08/MD-AIMF-3Q16-Executive-Summary.pdf

Kearns, Michael and Yuriy Nevmyvaka. “Machine Learning for Market Microstructure and High Frequency Trading.” (2013). https://www.cis.upenn.edu/~mkearns/papers/KearnsNevmyvakaHFTRiskBooks.pdf

Kearns, Michael, Yi Feng and Yuriy Nevmyvaka. “Reinforcement Learning for Optimized Trade Execution.” *Proceedings of the 23rd international conference on Machine learning* (June 2006): 673-680. https://www.seas.upenn.edu/~mkearns/papers/rlexec.pdf

Kercheval, Alec N. and Yuan Zhang. “Modelling high-frequency limit order book dynamics with support vector machines.” *Quantitative Finance*, Vol. 15, Issue 8, (October

2013):1315-1329, DOI: 10.1080/14697688.2015.1032546
https://www.math.fsu.edu/~aluffi/archive/paper462.pdf

Kolanovic, Marko and Rajesh T. Krishnamach. "Big Data and AI Strategies: Machine Learning and Alternative Data Approach to Investing." J.P. Morgan Global Quantitative & Derivatives Strategy (May 18 2017).
https://faculty.sites.uci.edu/pjorion/files/2018/05/JPM-2017-MachineLearningInvestments.pdf

Krzyk, Kamil. "Coding Deep Learning For Beginners." *Medium*, (July 25 2018).
https://towardsdatascience.com/coding-deep-learning-for-beginners-types-of-machine-learning-b9e651e1ed9d

Kumar, Chethan. "Artificial Intelligence: Definition, Types, Examples, Technologies." *Medium*, (August 31 2018).
https://medium.com/@chethankumargn/artificial-intelligence-definition-types-examples-technologies-962ea75c7b9b

Li, Yuxi. "Deep Reinforcement Learning." *arXiv:1810.06339* (Oct 15 2018).
https://arxiv.org/pdf/1810.06339.pdf

Li, Yuxi. "Reinforcement Learning Applications." *Medium* (October 15 2018).
https://medium.com/@yuxili/rl-applications-73ef685c07eb

Li, Yuxi. "Resources for Deep Reinforcement Learning." *Medium*, (December 28 2018).
https://medium.com/@yuxili/resources-for-deep-reinforcement-learning-a5fdf2dc730f

Liang, Percy. "Markov Decision Processes I." *CS221: Artificial Intelligence: Principles and Techniques* (Class Lecture 7, Stanford University, Autumn 2018).
http://web.stanford.edu/class/cs221/lectures/mdp1.pdf

Liang, Percy. "Markov Decision Processes II." *CS221: Artificial Intelligence: Principles and Techniques* (Class Lecture 8, Stanford University, Autumn 2018).
http://web.stanford.edu/class/cs221/lectures/mdp2.pdf

Lillicrap, Timothy, Jonathan J. Hunt, Alexander Pritzel, Nicolas Heess, Tom Erez, Yuval Tassa, David Silver and Daan Wierstra. "Continuous Control With Deep Reinforcement Learning." *Proceedings from the International Conference on Learning representations (ICLR)*, (2016).
https://arxiv.org/pdf/1509.02971.pdf

Loder, Asjylyn. "Goldman Rolls Out New ETFs Focused On Artificial Intelligence." *The Wall Street Journal*, (March 2019).

https://www.wsj.com/articles/goldman-rolls-out-new-etfs-focused-on-artificial-intelligence-11551978432

Lohr, Steve. "IBM Is Counting on Its Bet on Watson, and Paying Big Money for It." *The New York Times*, (October 2016). https://www.nytimes.com/2016/10/17/technology/ibm-is-counting-on-its-bet-on-watson-and-paying-big-money-for-it.html?emc=edit_th_20161017&nl=todaysheadlines&nlid=62816440

López de Prado, Marcos. "The 10 Reasons Most Machine Learning Funds Fail." *The Journal of Portfolio Management* Vol. 44, Issue 6 (2018): 120-133. https://jpm.iijournals.com/content/iijpormgmt/44/6/120.full.pdf

Mandel, Travis, Yun-En Liu, Emma Brunskill, and Zoran Popovic. "Offline Evaluation of Online Reinforcement Learning Algorithms." *In Proceedings of the 13th AAAI Conference on Artificial Intelligence*, AAAI Press 1926-1933 (February 2016). http://grail.cs.washington.edu/projects/nonstationaryeval/nonstationaryevalExtended.pdf

McCulloch, Warren and Walter Pitts. "A Logical Calculus of Ideas Immanent in Nervous Activity." *Bulletin of Mathematical Biophysics*. Vol. 5, Issue 4 (1943): 115–133. doi:10.1007/BF02478259

Mnih, Volodymyr, Koray Kavukcuoglu, David Silver, Andrei A. Rusu, Joel Veness, Marc G. Bellemare, Alex Graves, Martin Riedmiller, Andreas K. Fidjeland, Georg Ostrovski, et al. "Human level control through deep reinforcement learning." *Nature*, 518(7540):529–533, 2015. https://www.nature.com/articles/nature14236

Montantes, James. "Reinforcement Learning: From Grid World to Self-Driving Cars." *Medium* (January 31 2019). https://towardsdatascience.com/reinforcement-learning-from-grid-world-to-self-driving-cars-52bd3e647bc4

Motif. "The next wave of innovation is here." Accessed March 21 2019. https://www.motif.com/products/next-wave-portfolio

MSV, Janakiram. "In The Era Of Artificial Intelligence, GPUs Are The New CPUs." *Forbes* (August 7 2017). https://www.forbes.com/sites/janakirammsv/2017/08/07/in-the-era-of-artificial-intelligence-gpus-are-the-new-cpus/#54efc3e95d16

Nagabandi, Anusha, Gregory Kahn, Ronald S. Fearing and Sergey Levine. "Neural Network Dynamics for Model-Based Deep Reinforcement Learning with Model-Free Fine-Tuning."

2018 IEEE International Conference on Robotics and Automation (ICRA),DOI: 10.1109/ICRA.2018.8463189, (May 25 2018). https://arxiv.org/pdf/1708.02596.pdf

Ng, Andrew. "Reinforcement Learning and Control." *CS229: Machine Learning* (Class Lecture Notes 17-20, Stanford University, Autumn 2018). http://cs229.stanford.edu/notes/cs229-notes12.pdf

Nicholls, Harry. "7 Popular Technical Indicators and How to Use Them to Increase Your Trading Profits." *The Medium*, (April 8 2018). https://medium.com/@harrynicholls/7-popular-technical-indicators-and-how-to-use-them-to-increase-your-trading-profits-7f13ffeb8d05

Noel, Kevin. "Application of Machine Learning to Systematic Strategies." *SSRN Electronic Journal* (June 16, 2016). https://ssrn.com/abstract=2837664

Osband, Ian, Benjamin Van Roy and Daniel Russo. "(More) efficient reinforcement learning via posterior sampling." *In Proceedings of the 26th International Conference on Neural Information Processing Systems,* Volume 2 (NIPS'13) (2013). https://papers.nips.cc/paper/5185-more-efficient-reinforcement-learning-via-posterior-sampling.pdf

QuantMinds 365. "The latest in LOXM and why we shouldn't be using single stock algos." *KNect365*, (May 2018). https://knect365.com/quantminds/article/10d3b420-fe65-4269-b1da-ab555a509958/the-latest-in-loxm-and-why-we-shouldnt-be-using-single-stock-algos?utm_source=LinkedIn&utm_medium=Social&utm_campaign=LinkedIn-to-blog&utm_content=LIB

Ritter, Gordon. "Machine Learning for Trading." *SSRN Electronic Journal* (August 2017). https://ssrn.com/abstract=3015609

Rummery, G.A. and M. Niranjan. "On-Line Q-Learning using Connectionist Systems." Technical Report, CUED/F-INFENG/TR 166, Cambridge University (1994). ftp://mi.eng.cam.ac.uk/pub/reports/auto-pdf/rummery_tr166.pdf

Schwartz, Anton. "A Reinforcement Learning Method for Maximizing Undiscounted Rewards." *ICML* (1993). http://ftp.cs.stanford.edu/cs/robotics/schwartz/ml93.ps.gz

Sepp, Artur. "Machine Learning for Volatility Trading." *QuantMinds Invest Summit*, Lisbon (Presentation Slides) (May 2018). https://ssrn.com/abstract=3186401

Skywatch. "4 Ways satellite imagery is changing how we invest." Skywatch (blog) (August 22, 2016). https://www.skywatch.co/blog/4-ways-satellite-imagery-is-changing-how-we-invest

Silver, David, Guy Lever, Nicolas Heess, Thomas Degris, Daan Wierstra, Martin Riedmiller. "Deterministic Policy Gradient Algorithms." Proceedings of the 31st ICML, Beijing, China, (2014). JMLR: W&CP volume 32. http://proceedings.mlr.press/v32/silver14.pdf

Sirignano, Justin. "Deep Learning for Limit Order Books." *eprint arXiv:1601.01987* (January 2016). https://arxiv.org/pdf/1601.01987.pdf

Statt, Nick. "The AI boom is happening all over the world, and it's accelerating quickly." *The Verge*, (December 2018). https://www.theverge.com/2018/12/12/18136929/artificial-intelligence-ai-index-report-2018-machine-learning-global-progress-research

Sutton, Richard S. and Andrew G. Barto. *Reinforcement Learning: An Introduction*. 2nd edition, Cambridge, Massachusetts: The MIT Press, A Bradford Book, (November 5 2018). http://incompleteideas.net/book/RLbook2018.pdf

Taaffeite. "Taaffeite Capital Management." taaffeitecm.com, Accessed March 22 2019.

Terekhova, Maria. "JPMorgan takes AI use to the next level." *Business Insider*, (August 2 2017). https://www.businessinsider.com/jpmorgan-takes-ai-use-to-the-next-level-2017-8

Tsantekidis, Avraam, Nikolaos Passalis, Anastasios Tefas, Juho Kanniainen, Moncef Gabbouj & Alexandros Iosifidis. "Forecasting Stock Prices from the Limit Order Book Using Convolutional Neural Networks." *IEEE 19th Conference on Business Informatics* (2017) DOI: 10.1109/CBI.2017.23. http://poseidon.csd.auth.gr/papers/PUBLISHED/CONFERENCE/pdf/2017/2017_CBI_CNNLOB.pdf

Two Sigma. "Approach." *twosigma.com*, Accessed March 22 2019. https://www.twosigma.com/about/approach/

Udacity. "Deep Reinforcement Learning." Nanodegree program. https://www.udacity.com/course/deep-reinforcement-learning-nanodegree--nd893

Vapnik, Vladmir N.. "Statistical Learning Theory." *John Wiley & Sons, Inc*., Toronto Canada, (1998). https://www.amazon.ca/Statistical-Learning-Theory-Vladimir-Vapnik/dp/0471030031

Watkins, Christopher and Peter Dayan. "Q-Learning." *Machine Learning*, 8(3):279–292. http://www.gatsby.ucl.ac.uk/~dayan/papers/cjch.pdf

Wawrzynski, Pawel. "Real-time reinforcement learning by sequential Actor-Critics and experience replay." *Neural networks*: the official journal of the International Neural Network

Society. 22. 1484-97 (June 2019). http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.904.8961&rep=rep1&type=pdf

Yang, Qing, Tingting Ye and Liangliang Zhang. "A General Framework of Optimal Investment." *SSRN* (February 2, 2019). https://ssrn.com/abstract=3136708

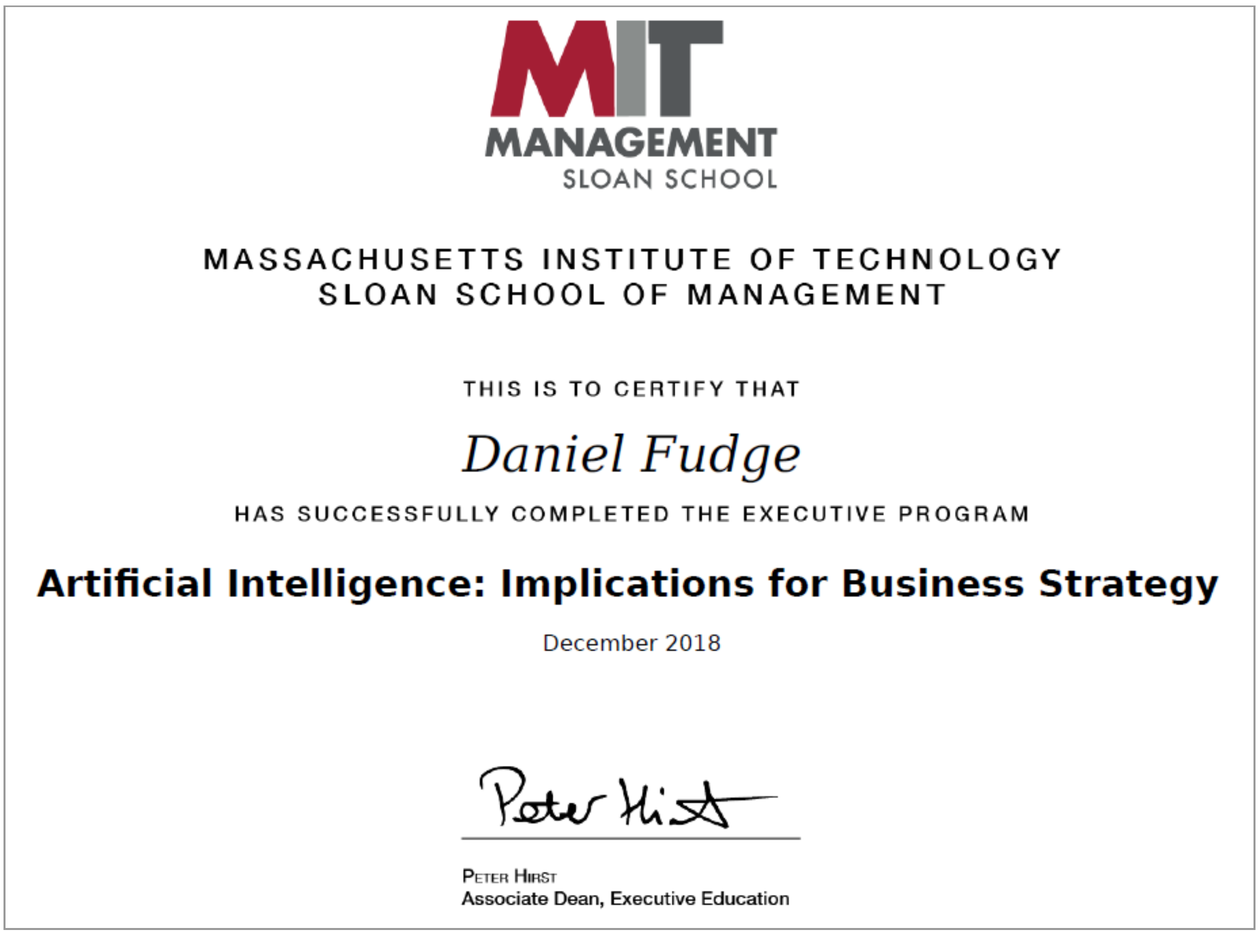
Appendix A - Related Training

A.1 MIT AI: Implications for Business Strategy

The Massachusetts Institute of Technology (MIT) Sloan School of Management and the MIT Computer Science & Artificial Intelligence Lab (CSAIL) delivered an online 6-week course call "Artificial Intelligence: Implications for Business Strategy."

https://executive.mit.edu/openenrollment/program/artificial-intelligence-implications-for-business-strategy-self-paced-online/#.XIHxbyhKiUk

I completed this course in December in anticipation for this research.


MIT MANAGEMENT SLOAN SCHOOL

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
SLOAN SCHOOL OF MANAGEMENT

THIS IS TO CERTIFY THAT

Daniel Fudge

HAS SUCCESSFULLY COMPLETED THE EXECUTIVE PROGRAM

Artificial Intelligence: Implications for Business Strategy

December 2018

Peter Hirst

PETER HIRST
Associate Dean, Executive Education

A.2 AWS Certified Developer - Associate

In December I also passed the AWS associate level developer certification. This was performed to increase my knowledge of cloud computing and how to implement a machine learning based portfolio management architecture.

https://aws.amazon.com/certification/certified-developer-associate/

aws CERTIFIED

Daniel Fudge

has successfully completed the AWS Certification
requirements and has achieved their:

AWS Certified Developer - Associate

Issue Date
December 22, 2018

Expiration Date
December 22, 2020

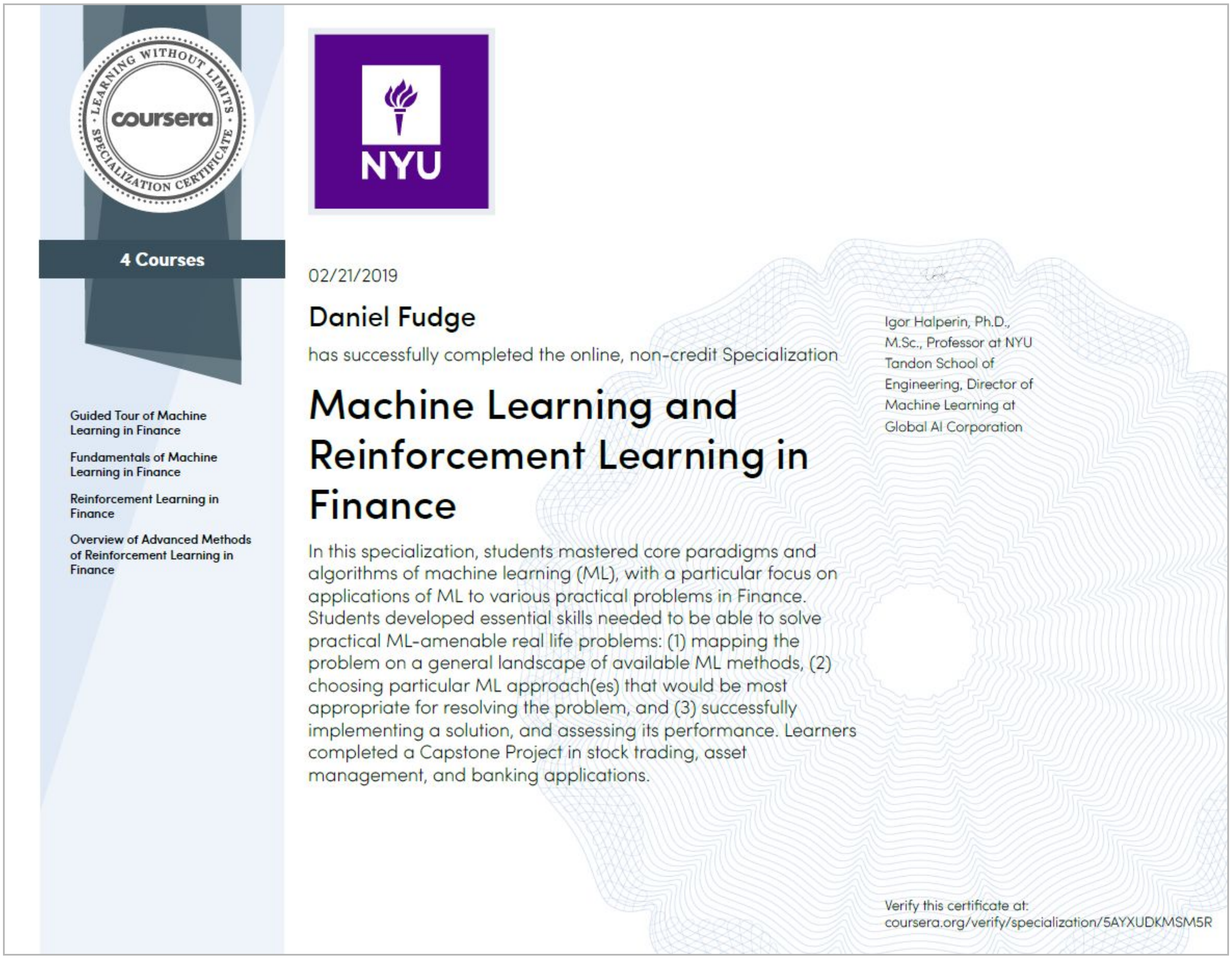
Maureen Lonergan
Director, Training and Certification

Validation Number JEJVMRPK12111KWC
Validate at: http://aws.amazon.com/verification

A.3 NYU Machine and Reinforcement Learning in Finance

This is a 5 month specialization consisting of 4 courses developed by Professor Igor Halperin and the New York University Tandon School of Engineering. I completed January and February of this term to support this research.

https://www.coursera.org/specializations/machine-learning-reinforcement-finance


LEARNING WITHOUT LIMITS · SPECIALIZATION CERTIFICATE

coursera

4 Courses

Guided Tour of Machine Learning in Finance

Fundamentals of Machine Learning in Finance

Reinforcement Learning in Finance

Overview of Advanced Methods of Reinforcement Learning in Finance

NYU

02/21/2019

Daniel Fudge

has successfully completed the online, non-credit Specialization

Machine Learning and Reinforcement Learning in Finance

In this specialization, students mastered core paradigms and algorithms of machine learning (ML), with a particular focus on applications of ML to various practical problems in Finance. Students developed essential skills needed to be able to solve practical ML-amenable real life problems: (1) mapping the problem on a general landscape of available ML methods, (2) choosing particular ML approach(es) that would be most appropriate for resolving the problem, and (3) successfully implementing a solution, and assessing its performance. Learners completed a Capstone Project in stock trading, asset management, and banking applications.

Igor Halperin, Ph.D., M.Sc., Professor at NYU Tandon School of Engineering, Director of Machine Learning at Global AI Corporation

Verify this certificate at:
coursera.org/verify/specialization/5AYXUDKMSM5R

Appendix B - Survey Material

B.1 Recruitment Text

The text below was sent to potential survey participants be sent via direct email and a LinkedIn Article linked here. Snowball sampling will then propagate the survey to the extended network.

Greetings,

My name is Daniel Fudge and I am a student in the Schulich School of Business concurrent MBA and Financial Engineering Graduate Diploma program (http://schulich.yorku.ca/programs/fnen/). With my supervisor Professor Yelena Larkin, I am conducting an independent study in potential applications of Machine Learning to Portfolio Management. As part of this research, I am trying to understand in what ways Machine Learning could be applied and used within financial institutions.

I am reaching out to you to ask if you would take the following 5-minute online survey. It is designed to anonymously collect information about Machine Learning applications in general, with an additional focus on Portfolio Management. No information identifying you or your organization will be requested or recorded.

The survey link is: Machine-Learning-Survey

A quick set of definitions including "Machine Learning" can be found here and a more informal description with some context can be found here. For an excellent introduction to machine learning, I recommend Andrew Ng's "Machine Learning" course from Stanford delivered by Coursera, which can be found here.

If you wish to learn more about this research or to obtain a copy of the final report, please feel free to contact Daniel Fudge at dfudge17@schulich.yorku.ca.

Best regards,

Daniel Fudge
P.Eng., M.ASc., MBA (in work)

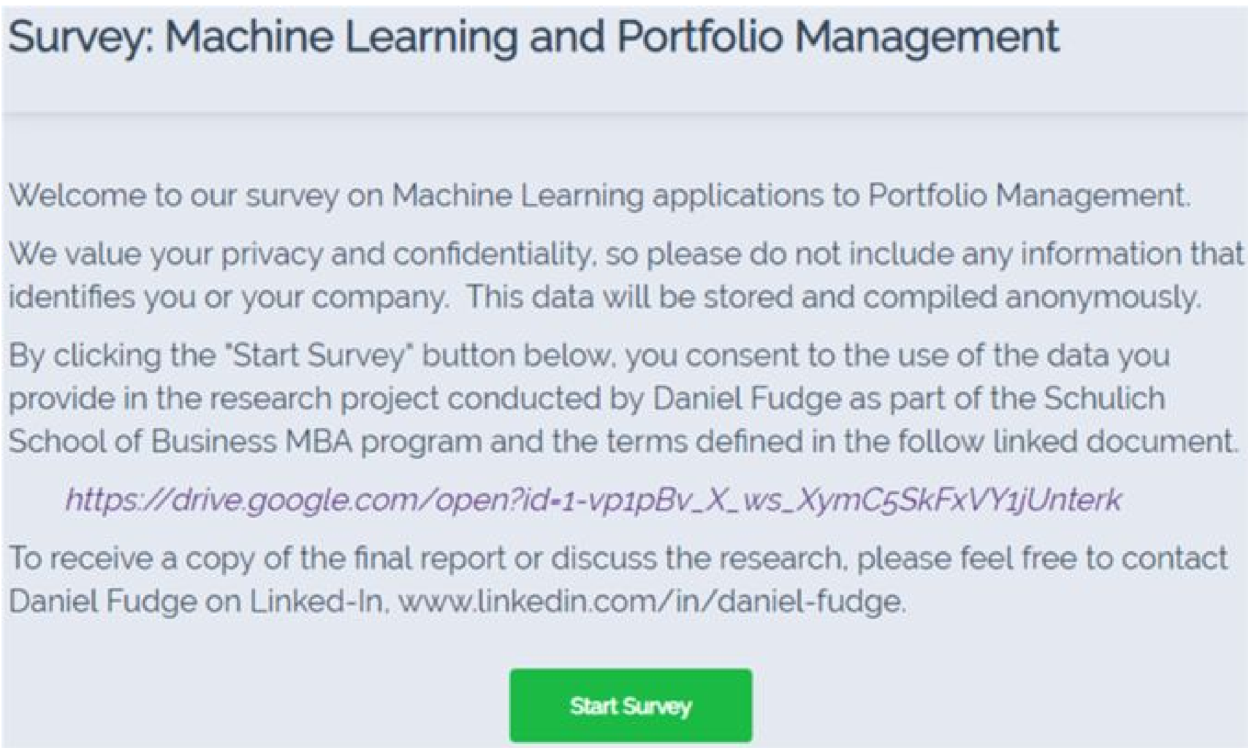
B.2 Survey Screenshots

The following screenshots were taken from the online survey.

Its link is https://app.qpointsurvey.com/s/dlMa2kEn8AHOzG3c8DIN9cChwk7iqUmPhNMYS0vlJFQ .

B.2.1 Welcome screen

Note the 2nd paragraph informs the participant of the anonymity of the survey, the 3rd contains the consent information and the 4th provides them a means to obtain more information.



B.2.2 Content Screen 1 of 3

Survey: Machine Learning and Portfolio Management

Please check the following items that you are aware of within your company?

- [ ] Application of Machine Learning in any department
- [ ] Research into Machine Learning by any department
- [ ] Application of Machine Learning for Portfolio Management
- [ ] Research into Machine Learning for Portfolio Management

Please check the following items that you are aware of within other financial institutions?

- [ ] Application of Machine Learning in any department
- [ ] Research into Machine Learning by any department
- [ ] Application of machine learning for Portfolio Management
- [ ] Research into Machine Learning for Portfolio Management

Next

B.2.3 Content Screen 2 of 3

Survey: Machine Learning and Portfolio Management

When do you think Machine Learning will impact Portfolio Management ?

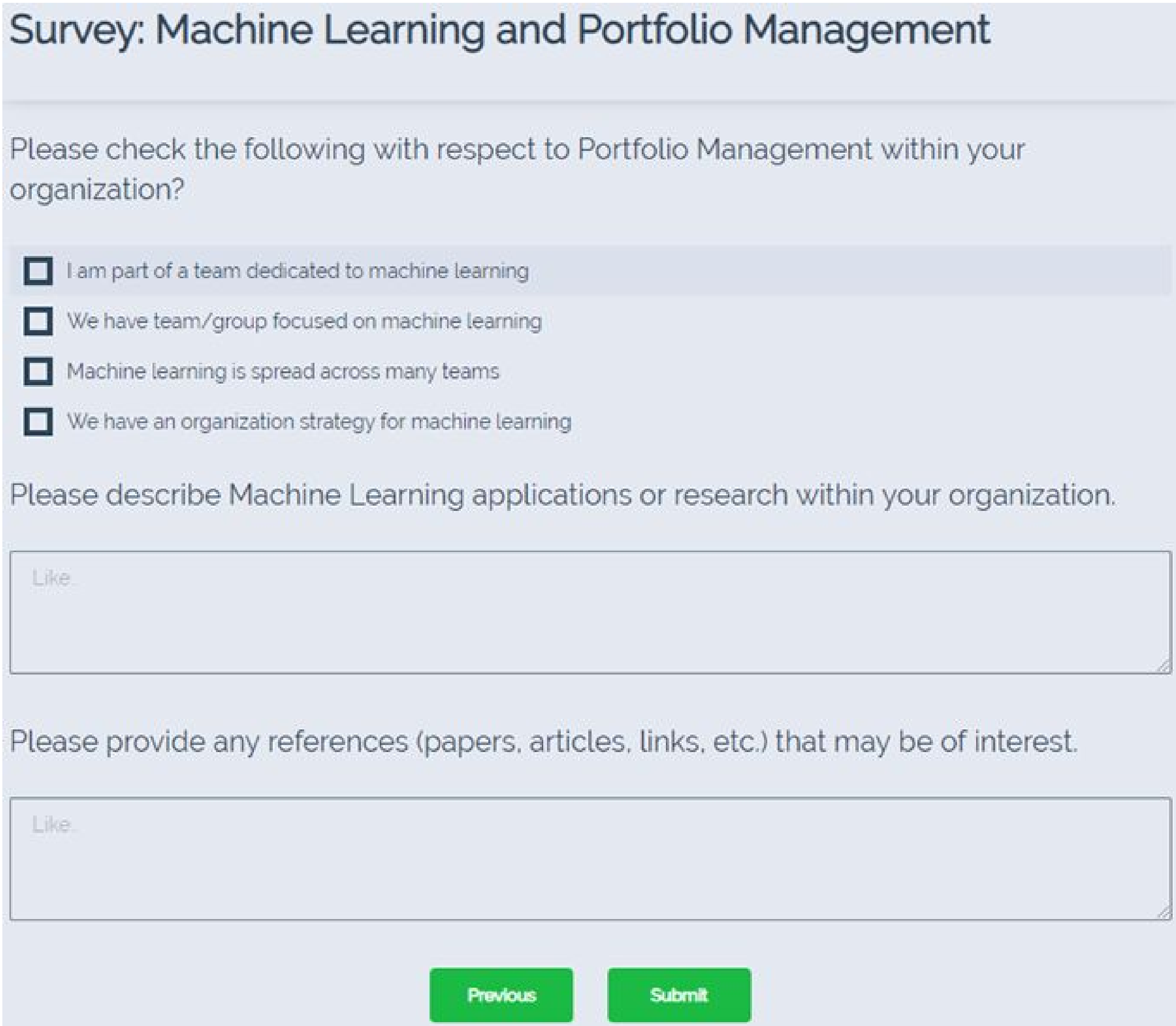
- It's making an impact today!!
- Within the next year
- Within the next decade
- Decades away
- Never

What is your organization's level of interest in Machine Learning for Portfolio Management?

- Never even discussed
- Occasionally discussed but no firm plans in place
- A research plan is being developed
- We are actively researching the capabilities
- We are actively applying the capabilities

Previous Next

B.2.4 Content Screen 3 of 3


Survey: Machine Learning and Portfolio Management

Please check the following with respect to Portfolio Management within your organization?

- [ ] I am part of a team dedicated to machine learning
- [ ] We have team/group focused on machine learning
- [ ] Machine learning is spread across many teams
- [ ] We have an organization strategy for machine learning

Please describe Machine Learning applications or research within your organization.

Like..

Please provide any references (papers, articles, links, etc.) that may be of interest.

Like..

Previous Submit

B.3 TCPS 2: CORE Certification

PANEL ON RESEARCH ETHICS
Navigating the ethics of human research

TCPS 2: CORE

Certificate of Completion

This document certifies that

Daniel Fudge

has completed the Tri-Council Policy Statement:
Ethical Conduct for Research Involving Humans
Course on Research Ethics (TCPS 2: CORE)

Date of Issue: **2 February, 2019**