1 Abstract

This project explores the opportunity of applying q-learning to create trading strategies for the financial markets. The process of the q-learning Problems and challenges are identified and discussed. Finally, a recommendation for future research is suggested.

2 Introduction

The possibilities of the applications of artificial intelligence have never been more current than now. Even though artificial intelligence is not a new phenomenon, it has been hard to apply because it often demands so much data and computation. As computing capacity increases, so does the feasible applications for AI and machine learning.

A field that has a history of data analysis to support decision making is the financial markets, and many actors are currently algorithms of different types to support their trading. The rationale that is often said being behind using an algorithms to either assist with or take investment decision is that the data does not lie, and that a computer will not get its feelings involved so that it cloud its judgment. Getting judgment clouded by emotions is possible and often though of as common among human investors, with behaviours such as animal spirits.
This report will investigate the possibility of applying the model free method of Q-learning, to train an algorithm for trading in equities. Depending on the desired strategy, for example what risk level to have, different value functions could be designed, e.g. Sharpe ratio, rate of return, or excess rate of return (often called alpha).

3 Input data

In order to train the algorithm, a large data set is needed. This data should include stock prices, and also fundamental data about the companies, such as profits, growth, ownership structure etcetera. This sort of data is available through for example FactSet. The more variables in the data set, the more possibilities of detecting relations that are currently unobserved and not exploited in the market. But more data will also make computation more demanding.

According to basic theories of supply and demand, the price of a good changes as supply or demand changes. This implies that when demand for an equity increases, so does the price of that equity, and vice versa for increased supply. Thus, an assumption that has to be made for this process, is that the trading algorithm is small enough is small enough to not move the market as it makes its trades. This means that for example that the price of a stock would not change if we choose to purchase some of them, which could be the case if we traded very large quantities. The reason for this assumption is that if the algorithm’s behaviour is expected to change the prices in the market, then it would not be possible to use historical data for the simulations in q-learning.

4 Process in the program

The process the program will utilise is a form of active reinforcement learning called Q-learning. Q-learning uses off-policy learning. The method works as you start with all Q-values equal to zero, and by exploration you get information which makes it possible to
update the Q-values according to the equation below:

\[ Q(s, a) < -(1 - \alpha)Q(s, a) + (\alpha)[sample] \]

Where:

\[ [sample] = R(s, a, s') + \gamma \max_{a'}Q(s', a') \]

Where the states contain the information in the used dataset such as stock prices and information about the profit of the companies listed on the stock market. The actions are defined as which equity to trade, how many (including if buying or selling), and at what price to trade.

5 Output

The output will be the optimal actions to take. From the estimated q-values we will see what action yields the highest utility, whichever way we choose to measure utility. As the actions are defined as what to trade at what price, the algorithm will return a recommendation for trading.

6 Potential problems

A big challenge for this project would be to choose an appropriate learning rate, as there is a tradeoff between exploration and exploitation. Another major challenge would be to have enough computational capacity to make the computations needed, in combination with trying to reduce needed computations to even make it feasible. One way of reducing the problem is reducing the number of states or actions available, this could be done by for example only looking at a subsection of companies on the stock market, or by reducing the number of attributes included in the states.

Sometimes, very unlikely so called "black swan" events occur in the market, which would
disrupt the effectiveness of the strategy. In addition to unlikely events, a crisis such as the
great recession of the late 00’s or the IT-bubble of the early 2000’s could cause a change
point in the relationship that the learning process has detected, possibly making it useless
after that crisis.

As financial markets work, for this algorithm to be able to yield utility, it is necessary
that not to many others utilise the same algorithm. This corresponds to the assumption
that our trading volumes should be small enough to not move the market, if many traders
used the same strategy, it would impact the valuations in the market.

7 Conclusion

The increase of computing capacity creates many great opportunities to implement machine
learning and artificial intelligence to the financial markets. However, there are many chal-
lenges. Some firms, such as Sentient use neural networks in combination with reinforcement
learning to produce their algorithms.

The use of Q-learning is however very demanding in terms of computational power, and
may thus not be the most useful or utilised method of implementing artificial intelligence
on the financial markets. Other methods for machine learning, such as neural networks
or MDP’s could also be used for the purpose of creating trading strategies, and should be
investigated further.

8 References

Satario, Adam, "Silicon Valley hedge fund takes on Wall Street with AI trader",
https://www.bloomberg.com/news/articles/2017-02-06/silicon-valley-hedge-fund-taking-on-wall-
street-with-ai-trader

Jiang, Zhengyao et al, "A deep reinforcement learning framework for the financial port-

