Automated Trading using Q Learning

Domain Background:

Automated trading algorithms have made their foray into the financial world. Over the years, the world of trading has evolved to leverage technology to create a differentiated advantage. In the recent past, machine learning has emerged as powerful weapon for trading firms to generate alpha and reduce costs. These tools have expanded across the domains of trading strategy, trade execution, settlement, compliance and operations.

Problem Statement:

The goal is to create an automated trading agent which will use Q learning (Reinforcement Learning) to predict trades for the Indian Nifty index, based on historical data and technical indicators (moving average, RSI, MACD).

Datasets and Inputs:

Historical data for Nifty index and volatility (India VIX) will be taken from the following links:

https://www.nseindia.com/products/content/equities/indices/historical_index_data.htm

https://www.nseindia.com/products/content/equities/indices/historical_vix.htm

Historical data for technical indicators will be taken from the following links:

http://www.traderscockpit.com/?pageView=macd-indicator-macd-chart&type=macd&symbol=NIFTY

http://www.traderscockpit.com/?pageView=rsi-indicator-rsi-chart&type=rsi&symbol=NIFTY

Solution Statement:

The states will be the technical indicators. The actions will be 'buy', 'sell', 'hold'.

The rewards will be the risk adjusted returns based on the above actions.

From the in-sample data, the agent will generate the Q-table consisting of all state-action pairs based on risk-adjusted returns (rewards). Based on this Q table, it will try to predict trades for the out-of-sample data.

Benchmark Model:

The return generated by the trading agent will be compared with the Nifty return for the same time period. The risk free rate for a 1-year bond would be considered 6.90%.

Evaluation Metrics:

The performance of the trading agent will be measured by the Sharpe ratio. Sharpe ratio measures the risk-adjusted return of a portfolio using the following formula:

Sharpe ratio = (return – risk free return)/volatility. So the reward in the Q learning algorithm will be equal to the Sharpe ratio.

Project Design Outline:

The NIFTY 50 is the flagship index on the National Stock Exchange of India Ltd. (NSE). The Index tracks the behavior of a portfolio of blue chip companies, the largest and most liquid Indian securities. It includes 50 of the approximately 1600 companies listed on the NSE, captures approximately 65% of its float-adjusted market capitalization and is a true reflection of the Indian stock market.

Buying 1 unit of Nifty would mean buying the entire portfolio of the 50 companies based on the weightages assigned in the index.

Since Q learning uses discrete states, the technical indicators (Moving Average, Volatility, RSI, MACD) will be converted to discrete values based on certain ranges. There will be 3 valid actions ('buy', 'sell', 'hold'). The agent can only buy or sell 1 unit at a time. Buying or selling more than 1 unit will involve a lot of complexity which I will not include in this project.

The in-sample period will be 1 year of Nifty index data (from 1st January 2017 to 31st December 2017). The out-of-sample will be 6 months of Nifty index data (from 1st January 2018 to 31st December 2018). I will only take the closing prices of Nifty to conduct the trades. Buying and selling will only happen at the closing price.

Initially, I will consider the entire asset to be invested in Nifty index. "Sell" would involve selling 1 unit of Nifty index and investing the proceeds in risk free 1-year bond. "Buy" would involve buying 1 unit of Nifty index using funds from risk free 1-year bond. So, the asset structure will be as follows:

[Funds in Nifty, Funds in Risk Free Bond, Total Funds]

Total Funds = Funds in Nifty + Funds in Risk Free Bond

A negative value for Funds in Risk Free Bond would mean that I borrowed that amount at a risk free rate. Similarly, a negative value for Funds in Nifty would mean I am short on Nifty and I have fully invested those funds in Risk Free Bond.

From the in sample period, the trading agent, by means of Q learning algorithm, will try to come up with the Q table consisting of all possible state-action pairs. Now it will use this Q table to predict trades for the out-of-sample period.

References:

http://houseofbots.com/news-detail/2830-1-does-the-future-of-trading-belong-to-artificial-intelligence

https://www.nseindia.com/content/indices/Method_Nifty_50.pdf