REINFORCEMENT LEARNING ON STOCK MARKET INVESTING

Stocks are one of the most dynamic environments which makes it quite difficult to determine the future with accuracy. However, stochastically we may be able to optimize the overall return from the stocks. As Reinforcement Learning uses stochastic models (via Markov Decision Process) to optimize the overall return, we try to explore various RL models to get a good overall return.

Stock market may exhibit some trend and seasonality but can exhibit high degree of unpredictability due to various reasons. Few of the well know scenarios are listed below:

* Covid-19 pandemic
* Israel-Palestine war
* Russia-Ukraine war
* Diplomatic dispute between India and Canada

Even across stocks, we can see that small-cap stocks are highly volatile due to pump and dump strategies adopted by market manipulators. Large-cap (and mid-cap to some extend) are comparatively immune to market manipulation strategies.

To derive best **ROI** (**R**eturn **O**n **I**nvestment), it is crucial to figure of the best time to **BUY** and best time to **SELL**. This implicitly also mean that we must know how long should we hold the stock before selling it off.

To aid us in decision making about the stocks, we have **Technical Indicators** to our rescue. These are various time series models which can tell us about the best time to BUY/SELL our stocks. These are not 100% accurate models, but when used together can provide us some better insights about the stock and can help in maximizing the overall return.

Few of the commonly used technical indicators have been listed below:

* MACD ([What Is MACD? (investopedia.com)](https://www.investopedia.com/terms/m/macd.asp))
* RSI ([Relative Strength Index (RSI) Indicator Explained With Formula (investopedia.com)](https://www.investopedia.com/terms/r/rsi.asp))
* BOLLINGER BAND ([Bollinger Bands: What They Are, and What They Tell Investors (investopedia.com)](https://www.investopedia.com/terms/b/bollingerbands.asp))
* STOCHASTIC MOMENTUM

In tandem, people also use the support and resistance levels to figure out the best price to invest or exit from the market. Overview of various technical indicators can easily be found over internet.

In most of the resource which talks about the usage of technical indicators do specify to use multiple of them together to get a better insight into the market scenario.

Analysing stock market data has some limitations. We have access to only few information like the opening price, closing price, highest value, lowest value all corresponding to a day. We don’t have access to how the stock have fluctuated in a day. Based on the limited resource available, we try to formulate our Markov Decision Process (MDP) which can be used to train our reinforcement learning model.

The MDP (s,a,r,s’) can be defined as:

* s: current states
* a: actions which we take on the current state
* r: reward received when we take the action
* s’: next state reached due to the action taken

Current State:

We have the stock price for everyday but we cannot use it as a state because just stock price alone doesn’t give us any indication if the stock is good to buy or not. As trading indicators do provide some insight about when to buy or sell, we consider the trading indicators as the current state.

Actions:

We consider actions as [BUY, HOLD, SELL] represented by values [1,0,-1]

Reward:

We are considering the reward as next day closing – current day closing.

Next State:

We have the information of the technical indicators for next day as well. This is what we consider as next state.