Reinforcement Learning for Foreign Exchange Trading

Jiahao Zhang, Ruojing Song, Yihui Zhang

Introduction

The foreign exchange market is the largest and most liquid financial market in the world. With the development of the computation and machine learning techniques, we are curious about whether AI can help address problems in decision making based on given conditioning information of the foreign exchange market. In this project, we compared the performance of two reinforcement learning techniques, DQN and SARSA, with other reflex machine learning models in profiting from foreign exchange transactions.

Data

From Bloomberg, we collected:

- US Dollar to British pound (USD/GBP)
- US Dollar to Euro (USD/EUR)
- US Dollar to Japanese Yen (USD/JPY)
- More macro-economic and financial features

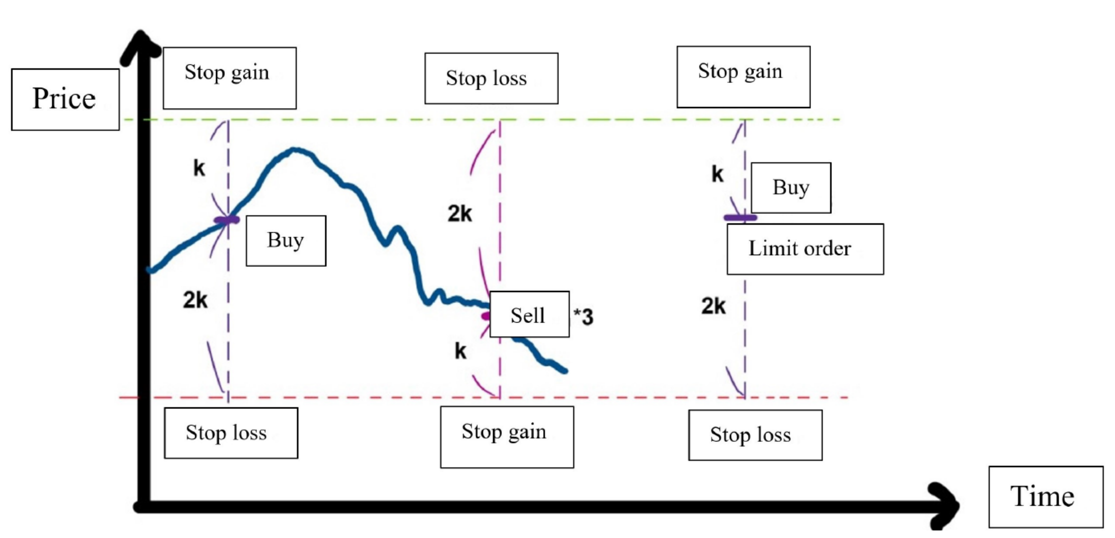
Approach / Methodology

We extensively tested different kinds of financial and machine learning techniques in the application of foreign exchange forecasting, including:

- Random Forest (Experiment 1: Reflex)
- Sure-Fire Strategy (Experiment 2: Reinforcement Learning)
- Deep Q-Learning (Experiment 2: Reinforcement Learning)
- Deep SARSA (Experiment 2: Reinforcement Learning)

Sure-Fire Trading Strategy

The main idea of Sure-Fire is by continuously raising the funds until the currency exchange rate either surpasses a upper stop-limit or falls below a lower stop-limit, the last transaction can recover all the previous losses plus a certain amount of profit.



Experiment 1: Reflex

Objective:
Maximize profit via long/short currency pairs within a period of time

Evaluation:
We invested with a fixed capital of 100 dollars and trade between the three currency pairs without shorting.
First, compute the predicted return of the currency pair at time t:

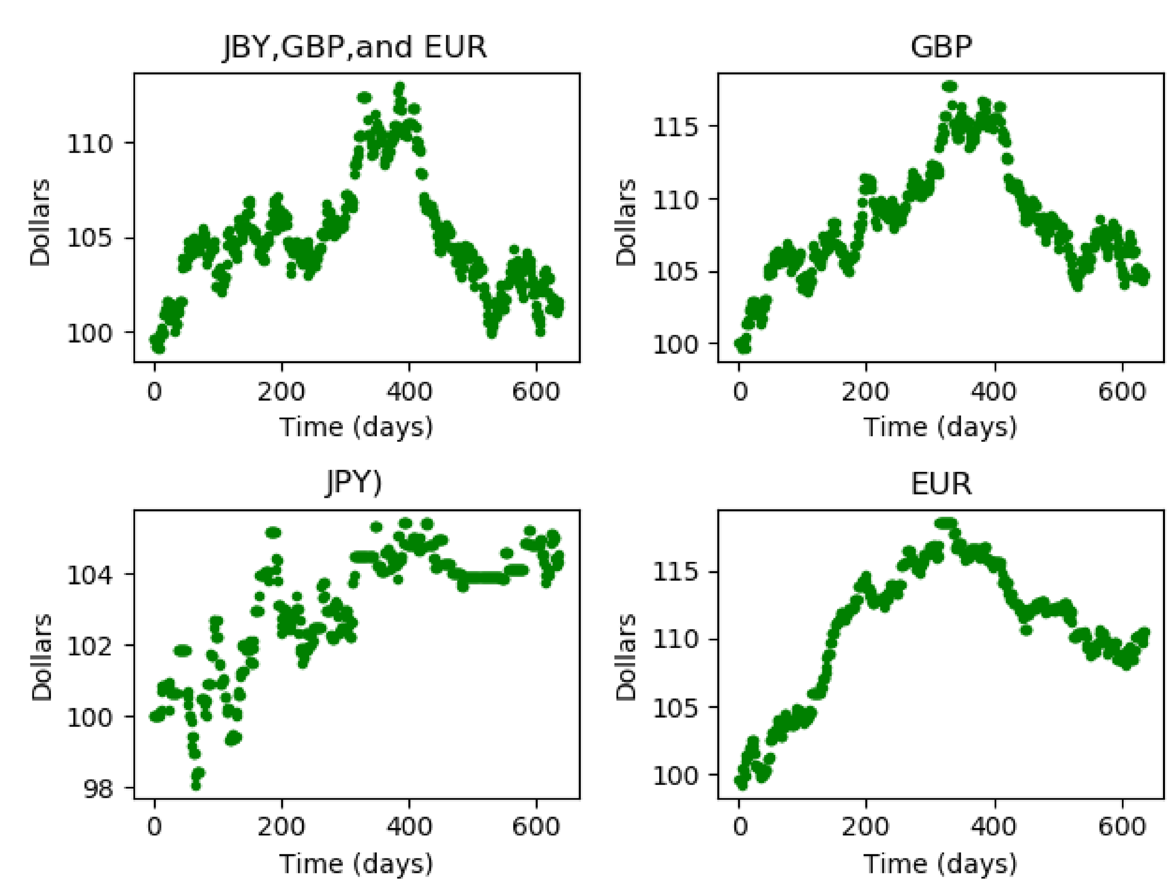
$$\text{return}_t = \frac{\text{ExchangeRate}_t}{\text{ExchangeRate}_{t-1}} \tag{1}$$

Second, if all return values are larger than 1, we do nothing; if there are some return values lower than 1, then we "switch" to the currency that has lowest exchange rate return with USD for one day and "switch" back.

Experiment 1 Result

Profit Simulator:

	GBP	JPY	EUR
Final	104.40	104.44	110.5
	GBP+JPY	JPY+EUR	GBP+EUR
Final	100.33	109.46	104.88
	GBP+JPY+EUR		
Final	101.30		



Our portfolio value peaks at about March 2018, but eventually drops. We investigated thoroughly about the reasons for the failure of our trading strategy after the point and one explanation is that the trade war between the U.S and China, which took place exactly at that time, had changed the regime of the foreign exchange rate dynamics.

Experiment 2: Reinforcement Learning

Deep Q-Learning:
Deep Q-Learning algorithm(DQN) is a modified deep reinforcement learning framework. With the same objective as Reflex model, DQN's input data consists of smoothed out time series financial data, and its fully-connect layers can take in input data and generates n by 1 vectors to update k value. In addition, every time an action is selected, the system provides a small probability (e) to explore unknown actions, where skip the current best function and execute new actions. The DQN's update function is as followed:

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha[r_{t+1} + \gamma \max_a Q(s_{t+1}, a) - Q(s_t, a_t)]$$

Deep SARSA:
D-SARSA resembles DQN's method with the main difference that it's on-policy. D-SARSA we use the same policy to generate the current action a_1 the next action a_{t+1}. Then policy's action selection is evaluated, and improved upon by improving the Q-function estimates. D-SARSA's update function is as followed:

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha[r_{t+1} + \gamma Q(s_{t+1}, a_{t+1}) - Q(s_t, a_t)]$$

Experiment 2 Result and Analysis

Both DQN and D-SARSA algorithms turn out to have the ability to make a net profit at the end of the test period by choosing a sequential group of actions. Each action is a combination of the three controllable variables in the Sure-Fire strategy, which are the upper limit of number of tradings (a_1), first BUY or SELL as entering the market (a_2), and the stop-gain, k (a_3).

However, foreign exchange market is highly volatile with uncertain patterns that are not predictable, it is difficult for RL to make decisions. Among the test results of three currency pairs, convergence on the net profit as the number of training episodes increases is only observed on USD/JPY pair. For the rest two, as the training episode approaches to the end, net profit fluctuates within a small range. Therefore, we use the average of the net profits in the last 20 episodes (out of 2000 episodes in the entire training period) as our **objective metric** for the evaluation of agent performance:

	JPY/USD	EUR/USD	GBP/USD
Test Period	08/16/2013 - 12/31/2018		
Unit Price at 08/16/2013	97.53	0.7503	0.6399
Net Profit per Unit at 12/31/2018	21.22	0.1080	0.1752

Other conclusions that are derived from the RL experiment:

- As Q-learning has higher per-sample variance than SARSA, in our case, DQN is much harder to converge than D-SARSA.
- The 15 features currently used as the input to the neural network cannot evaluate the Q-value very well. Since most of these features only represent the economical and financial situation in America, adding features such as crude oil price and gold price which are global financial indicators will improve the performance.
- Among the many tunable hyperparameters in the realm of DQN and D-SARSA, key hyperparameters that largely affect the performance are: discount factor, epsilon_decay_steps, and learning rate.