Literature Review

High Frequency Portfolio Optimization using Deep Reinforcement Learning

Group 10: Xiaoyu Chen, Zhiyuan Yao, Yuqi Liu, Zhixian Hu
Advisor: Dr. Ionut Florescu, Dr. Steve Yang

February 7th, 2019

STEVENS
INSTITUTE of TECHNOLOGY
THE INNOVATION UNIVERSITY®

Overview

- Deep Neural Network
- Reinforcement Learning
- Modeling on High Frequency Trading
- Inverse Reinforcement Learning

STEVENS
INSTITUTE of TECHNOLOGY
THE INNOVATION UNIVERSITY®

Deep Neural Network

- **Title:** Deep Learning for Spatio-Temporal Modeling- Dynamic Traffic Flows and High Frequency Trading
- **Author:** Matthew F. Dixon, Nicholas G. Polson, Vadim O. Sokolov
- **Source:** arXiv preprint arXiv:1705.09851
- **Year:** May 7, 2018

Deep Neural Network

- **Main contribution**
  - Algorithmic approach to predict short-term price movement
  - Classify book orders into directional mid-price movement {-1, 0, 1}
  - Demonstrate the ability of RNN to capture non-linear relationship

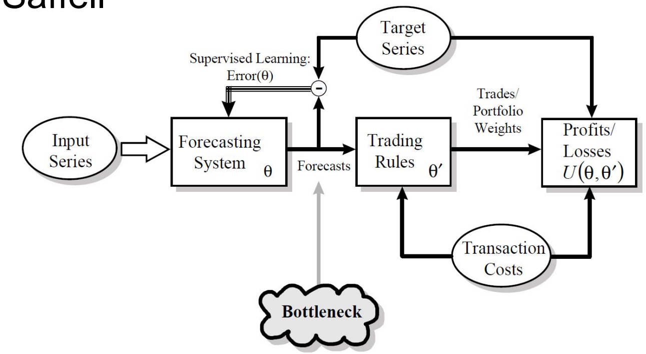
n_steps	$\hat{Y} = -1$	$\hat{Y} = 0$	$\hat{Y} = 1$
1	0.089	0.528	0.091
2	0.097	0.837	0.103
5	0.131	0.848	0.120
10	0.165	0.881	0.161
20	0.163	0.878	0.162
50	0.151	0.836	0.144
100	0.118	0.807	0.115

- **Main methods**
  - Recurrent Neural Network using SGD
  - Drop-out methods: RNN design

- **How help**
  - Feature construction: order-flow, liquidity imbalance..
  - Feature selection

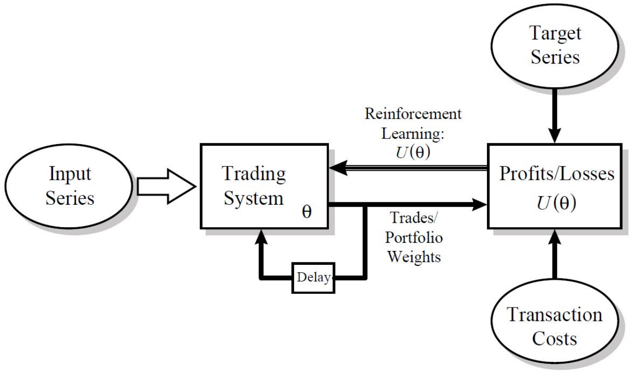
Reinforcement Learning in Trading

- **Title:** Performance Functions and Reinforcement Learning for Trading Systems and Portfolios
- **Author:** John Moody, Lizhong Wu, Yuansong Liao, Matthew Saffell
- **Source:** Journal of Forecasting, Volume 17, Pages 441-470
- **Year:** 1998



Key points:

- Recurrent Reinforcement Learning (RRL) trading system is better than trading system based on supervised learning and labelled data.
- Optimization via RRL
- The Sharpe Ratio, EMA Sharpe Ratio and Differential Sharpe Ratio
  - Advantages
  - Other performance ratios: Information ratios and Sterling Ratio
- Empirical results



Reinforcement Learning in Trading

- **Methods:**
  - Deriving the gradient of policy function by chain rule.
  - Using Exponential Moving estimates to estimate the first two moments of rate of return, then compute the the Sharpe ratio.
  - Using Taylor Expansion on EMA Sharpe ratio to obtain a increment expression which can be cheaply computed during online learning

- **Going Forward**
  - Action is only discretized into 3 statues (Buy, Sell and Hold). The situation becomes more complex when doing High Frequency Trading (HFT).
  - Based on the latest research on reinforcement learning, we can try policy optimization instead of policy gradient.
  - How to parametrize policy function is not specified.

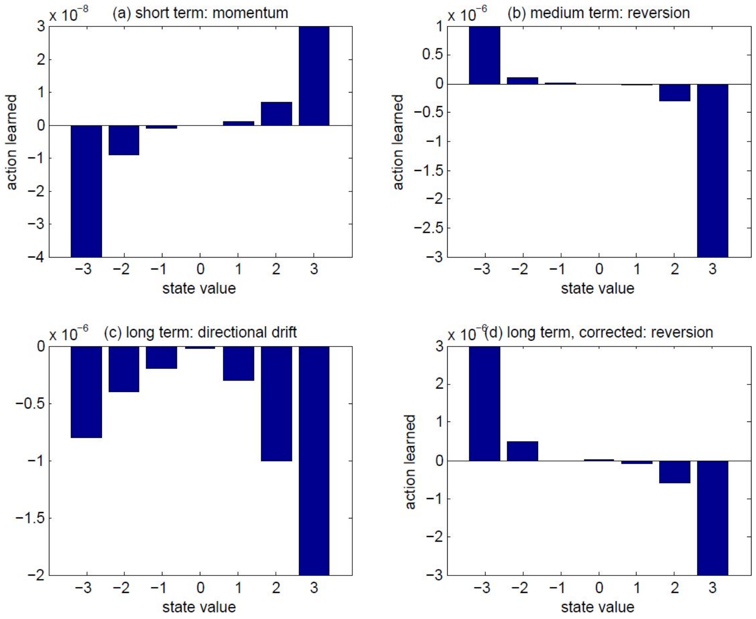
Modeling on High Frequency Trading

- **Title:** Machine Learning for Market Microstructure and High Frequency Trading
- **Author:** Michael Kearns, Yuriy Nevmyvaka
- **Source:** High Frequency Trading: New Realities for Traders, Markets, and Regulators
- **Year:** 2013

Modeling on High Frequency Trading

- Optimal Trade Execution & Price Movement
  - Reinforcement Learning
    - State: (v, t); additional features
    - Reward function: trading cost
  - Value Functions
    - Holding period: momentum/reversion
    - Longer horizon: less informative
- Give ideas about feature modeling and study further on strategy of time scale

Feature(s) Added	Reduction in Trading Cost
Bid-Ask Spread	7.97%
Bid-Ask Volume Imbalance	0.13%
Signed Transaction Volume	2.81%
Immediate Market Order Revenue	4.26%
Spread + signed Volume + Immediate Cost	12.85%

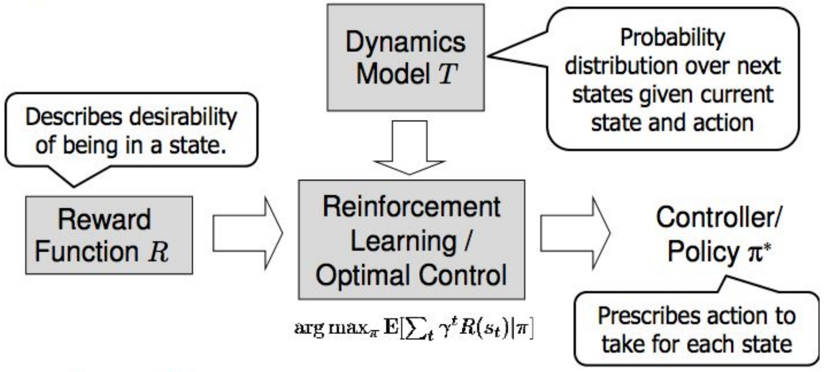


Inverse Reinforcement Learning

- **Title:** Behavior Based Learning in Identifying High Frequency Trading Strategies
- **Author:** Steve Yang, Mark Paddrik, Roy Hayes, Andrew Todd, Andrei Kirilenko, Peter Beling, and William Scherer
- **Source**: Computational Intelligence for Financial Engineering & Economics (CIFEr), 2012 IEEE Conference on (pp. 1-8). IEEE.
- **Year:** March, 2012

Inverse Reinforcement Learning

- **Obtain Reward Function & Optimal Policy**
  - MDP Model for limit order book
  - Inverse Reinforcement Learning
- **Going Forward**
  - Dealing with limit order book with MDP
  - Policy Optimization by IRL Model



Thank You

School of Business

Zhiyuan Yao | zyao9@stevens.edu

Zhixian Hu | zhu17@stevens.edu

Xiaoyu Chen | xchen79@stevens.edu

Yuqi Liu | yliu176@stevens.edu