Deep Portfolio Management
A summary of Deep Reinforcement Learning for the Financial Portfolio Management Problem (Jiang et. al.) by Shawn Anderson

The financial portfolio management problem is an optimization task in which we have some funds which we wish to invest in a portfolio of assets such that the growth of our funds is maximized. Furthermore, we break time into periods of length T. At the start of each period t, we redistribute our funds across the assets in order to optimize growth over that period of time. At the end of the period we can calculate our total portfolio value to see if our funds have increased and by how much. In this way, the problem can be modeled as a deep reinforcement learning problem where the reward for a period is the increase in portfolio value, or return, and the environment is the price movement history over the last n periods.

In this paper, the authors use cryptocurrency markets for testing their solution. In particular, all testing is done with poloniex.com, a very popular cryptocurrency exchange. Cryptocurrencies are digital tokens of value built with cryptographic tools such that ownership and trade can be recorded using a decentralized, distributed, algorithmic system. Bitcoin is the most popular example of a cryptocurrency. The authors use cryptocurrency markets for reasons of convenience, such as continuously open markets (24/7), and open API's for data collection and automated trading. Furthermore, cryptocurrency markets are a new and emerging phenomenon that offer many avenues of scientific exploration in their own right.

All experiments in the paper are setup in the following way: Bitcoin is used as the 'cash', that is, the price of all assets are measured relative to Bitcoin. In other words, the price of Bitcoin at all time steps is constant at 1. The portfolio vector is always of length 12, consisting of the cash, plus 11 non-cash assets. Assets are selected by highest trade volume to accomodate two market assumptions that the authors make. All periods are of length 30 seconds, ie. the portfolio vector is redistributed every 30 seconds. The action of the RL agent is the redistribution of the portfolio vector. The environment of the RL agent is a 3xmxn price tensor which represents the price movements of m assets over n periods of history. The 3 represents the number of features for a given asset over a period, specifically, closing price, highest price, and lowest price.

The authors propose a novel architecture concept which they call Ensemble of Identical Independent Evaluators (EIIE). A very small architecture network trains over each asset independently.
In this method, model parameters are tied (identical) for the evaluation of all assets. There are no fully connected layers, asset information remains independet through the whole network until the softmax output layer. The authors implement EIIE using three different architectures, CNN, RNN, and LSTM.

The authors run three back tests, each with a training period of roughly two years, and a testing period of exactly 50 days. They use three performace metrics, final portfolio value, sharpe-ratio (favors high return, penalises risk), and Maximum Drawdown (greatest portfolio loss sequence in test). For each back-test they report three vesions of their solution, CNN, RNN, and LSTM. They compare results with three benchmarks (Best Stock, Uniform Buy and Hold, and UCRP), as well as 12 other popular portfolio optimization algorithms. In all three back tests, the EIIE methods hold top scores for final portfolio value and sharpe ratio. In back test 3, the RNN EIIE sees returns of 47x in 50 days.

The relavence to this class is the following 1. Deep Learning applied to finance 2. The novel EIIE approach 3. A peak at deep RL 4. A modular learning system in which the architecture can be easily substituted between CNN, RNN, LSTM