Evaluation of Machine Learning Trading Strategies Using Recurrent Reinforcement Learning (RRL) [1]

Noteworthy Trading Characteristics	*Choice of Algorithm*
❖ Investment performance depends upon sequences of interdependent decisions (Path-dependent)	➔ **Recurrent** to account for effects of transaction costs, market impact, etc
❖ Presence of large amounts noise and nonstationarity in the finance datasets	➔ **Policy, instead of value function**, is used. A much simpler functional form is often adequate.
❖ Short-term performance can be immediately measured and financial results accrue over time	➔ **Direct reinforcement** to the policy

Trader function:

$$F_t = tanh(w^T x_t) \in \{1,\ 0,\ -1\},\ i.e.\ long,\ neutral,\ or\ short$$

where $x_t = [1,\ r_t,\ \ldots,\ r_{t-M},\ F_{t-1}]$,
$r_t = \ log\ price\ return\ = ln\frac{p_t}{p_{t-1}}$
$M\ = \ Number\ of\ time\ series\ inputs,\ i.e.\ length\ of\ return\ array$

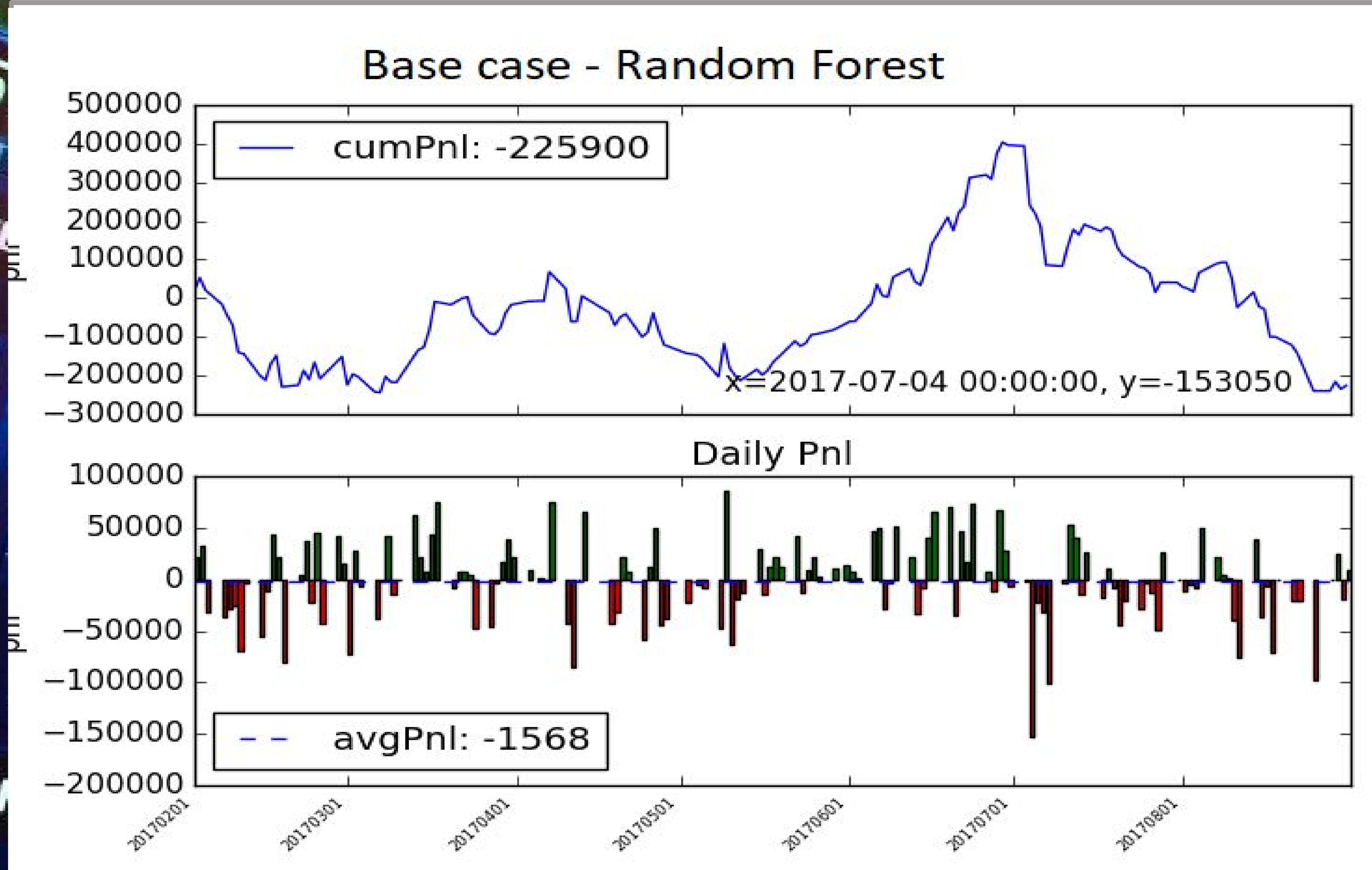
Maximizing return:

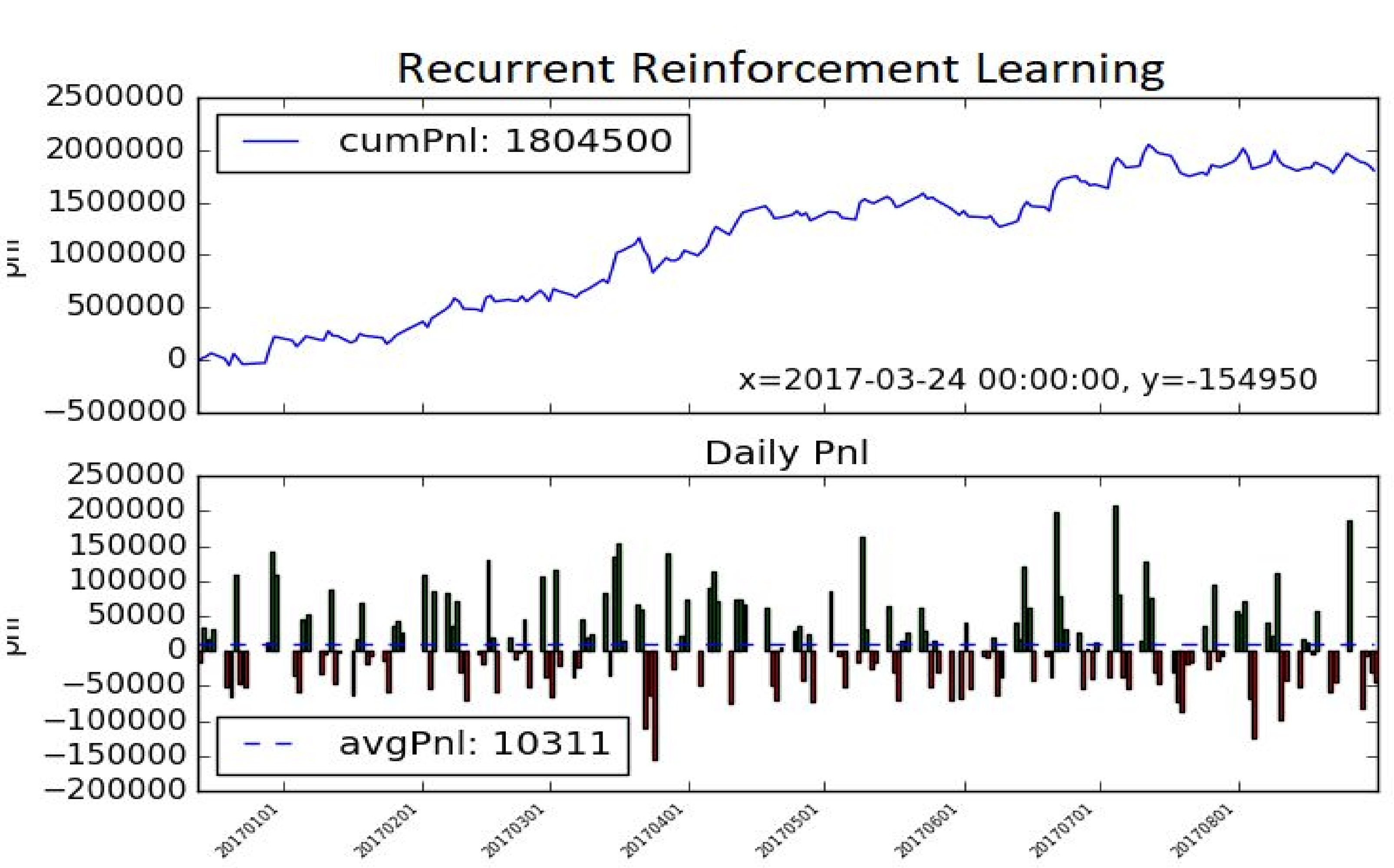
$$R_t = \mu\left(F_{t-1}v_t - \delta\ \left|F_t - F_{t-1}\right|\right)$$

where $\mu = maximum\ number\ of\ shares\ per\ transaction$
$\delta = transaction\ cost\ in\ bps$

Optimization variables: $\theta = \{v,\ v_i,\ w\}$

Preliminary results have verified that **RRL-traders outperforms more those based on standard supervised approaches**, using open high low close (OHLC) prices for the Hong Kong Hang Seng Index future.





Also using LSTMs to predict the next price (relative move). Initial results are mixed but tweaking parameters to see if we can improve results..

1. "Learning to Trade via Direct Reinforcement", John Moody and Matthew Saffell, 2001