
高级量化交易技术

闫涛
科技有限公司
北京 2021.05.08
{yt7589}@qq.com

第零篇深度学习

第 4 章回测平台

Abstract

在本章中我们将详细讲解基于强化学习环境的回测平台。

1 回测平台概述

在本章中，我们基于强化学习环境，创建一个策略的回测平台。

1.1 启动入口

下面我们来看启动的入口函数：

```
1 def run(self):
2     stock_symbol = 'sh600260'
3     model = self.build_model()
4     env = FmtsEnv(stock_symbol)
5     obs = env.reset()
6     done = False
7     action = env.action_space.sample()
8     idx = 0
9     while not done:
10        X = self.get_model_X(obs)
11        quotation_state = self.get_quotation_state(model, X)
12        if quotation_state == 0:
13            # 买入
14            action[0] = 0.5
15            action[1] = 1.0
16        elif quotation_state == 1:
17            # 卖出
18            action[0] = 1.5
19            action[1] = 1.0
20        else:
21            # 持有
22            action[0] = 2.5
23            action[1] = 0.5
24        obs, reward, done, info = env.step(action)
25        if env.current_step > 200:
26            done = True
27
28 def build_model(self):
29     '''
30     强化学习环境Reset中需要调用本函数，初始化模型
31     '''
32     cmd_args = self.parse_args()
33     print('command line args: {0};'.format(cmd_args))
34     batch_size = cmd_args.batch_size
35     NUM_CLS = 3
```

```

36     cmd_args.embedding_size = 5
37     seq_length = 11
38     cmd_args.depth = 6
39     cmd_args.num_heads = 8
40     model = FmtsTransformer(emb=cmd_args.embedding_size, heads=
cmd_args.num_heads, depth=cmd_args.depth, \
41         seq_length=seq_length, num_tokens=cmd_args.
vocab_size, num_classes=NUM_CLS, \
42         max_pool=cmd_args.max_pool)
43     model.to(self.device)
44     e, model_dict, optimizer_dict = self.load_ckpt(self.ckpt_file)
45     model.load_state_dict(model_dict)
46     return model

```

Listing 1: 回测平台入口

代码解读如下所示：

- 第 3 行：创建模型，具体代码见第 28 行；
- 第 4 行：创建强化学习环境，具体见下一节；
- 第 5 行：重置强化学习环境，并获得系统初始观察值；
- 第 6 行：将是否结束回测置为否；
- 第 7 行：从行动空间中随机抽样出一个行动，行动是一个二维数组，第 1 个元素代表操作，在 $[0,1]$ 之间为买入，在 $(1,2]$ 之间为卖出， $(2,+\infty)$ 时为持有，第二个元素为操作的百分比；
- 第 9 行：如果结束回测标志不为真则一直循环；
- 第 10 行：根据环境的观察求出用于模型的输入信号；
- 第 11 行：通过运行模型，确定当前时刻的市场行情状态：0-上涨、1-下跌、2-震荡；
- 第 12~15 行：当处于上涨行情时，如果还有现金，则执行买入操作，`action[0]` 取小于 1 的数，`action[1]` 指定使用当前现金的百分比；
- 第 16~19 行：当处于下跌行情时，如果持有股票，则执行卖出操作，`action[0]` 取 1 至 2 之间的数，`action[1]` 指定使用当前持股数的百分比；
- 第 20~23 行：当处于震荡行情时，`action[0]` 取大于 2 的数，`action[1]` 的值被忽略；
- 第 24 行：将行动传给环境执行，环境会返回：`obs`-当前的状态（执行完行动后）、`reward`-该行动获得的奖励、`done`-是否结束回测标志、`info`-其他附加信息；

1.2 强化学习环境

2 总结

