

基于 LSTM 股指预测的日内交易策略报告

一、策略一 ctaLSTM_V1 (对应预测模型 LSTM_V3.0)

1.策略内容：

利用 LSTM 模型完成对诸多指标的预测，就开始着手写交易策略了。目前的策略，只用到了当前时刻后 1-5 分钟的涨跌分类。该策略主要包含以下部分：

(1) 获取当前的 bar 数据，并添加上述的 34 个特征。归一化，生成时间序列，使格式与训练时用的数据集格式完全相同。

(2) 加载 LSTM 网络模型，获得后 1-5 分钟的分类结果。分类中，‘5’表示大涨，‘4’表示小涨，‘3’表示平稳，‘2’表示小跌，‘1’表示大跌。

代码如下：

```
# 将当前数据代入模型，获得1-5分钟的分类类别
# 5: 大涨 4: 小涨 3: 平稳 2: 小跌 1: 大跌
pred_1 = sess_1.run(pred, feed_dict={xtr: seq_data})
predict_1 = 5 - pred_1.argmax()
pred_2 = sess_2.run(pred, feed_dict={xtr: seq_data})
predict_2 = 5 - pred_2.argmax()
pred_3 = sess_3.run(pred, feed_dict={xtr: seq_data})
predict_3 = 5 - pred_3.argmax()
pred_4 = sess_4.run(pred, feed_dict={xtr: seq_data})
predict_4 = 5 - pred_4.argmax()
pred_5 = sess_5.run(pred, feed_dict={xtr: seq_data})
predict_5 = 5 - pred_5.argmax()

return predict_1, predict_2, predict_3, predict_4, predict_5
```

(3) 对 1-5 分钟的预测值赋予不同权重，时间越短权值越高，得到一个综合得分。

```
# 计算加权得分
self.score = 0.4 * self.predict_1 + 0.2 * self.predict_2 + \
    0.2 * self.predict_3 + 0.1 * self.predict_4 + \
    0.1 * self.predict_5
```

(4) 根据得分判断是否作出买卖请求。当得分大于一个设定的最高阈值时，表示后 1-5 分钟基本属于大涨趋势，可以买多；当得分小于一个设定的最低阈值时，表示后 1-5 分钟基本属于大跌趋势，可以买空。

```
if self.score >= self.buy_raise_score and self.zhisun_label is False:
    self.buy(self.askprice1, 1)
    self.rec_price = self.askprice1 # 记录当前的价格作为比较价格
    self.pos_rec += 1
    self.writeCtaLog(u'buy!' + str(self.rec_price) + u'pos_rec' + str(self.pos_rec))
    self.trade_records.append([bar.datetime, self.price, u'buy'])
    self.init_price = self.askprice1

if self.score <= self.buy_down_score and self.zhisun_label is False:
    self.short(self.bidprice1, 1)
    self.rec_price = self.bidprice1
    self.init_price = self.bidprice1
    self.pos_rec -= 1
    self.writeCtaLog(u'short!' + str(self.rec_price) + u'pos_rec' + str(self.pos_rec))
    self.trade_records.append([bar.datetime, self.price, u'short'])
```

(5) 止盈止损策略。同样的，设置一个止盈系数和一个止损系数。手中持有仓时，当前买一价（卖一价）已经超过了止盈（止损）所限定的最大范围，则平仓。

```

def long_pos_sell(self, long_pos):##多仓的平仓考虑。

    if long_pos.bidPrice1 < (1 - self.zhisun)*self.rec_price:##止损策略
        self.sell(self.bidprice1, 1)
        self.pos_rec -= 1
        self.writeCtaLog(u'多仓'+u'止损价'+str((1 - self.zhisun) * self.rec_price)+u'当前价'+str(self.price))
        self.zhisun_label = True
        self.zhisun_bar = 0
        self.trade_records.append([long_pos.datetime, self.price, u'sell'])

    if long_pos.bidPrice1 > self.rec_price*(1+self.zhiying):##止盈策略
        self.sell(self.bidprice1, 1)
        self.pos_rec -= 1

        self.writeCtaLog(u'多仓'+u'止盈价'+str(self.rec_price*(1+self.zhiying))+u'当前价'+str(self.price))
        self.trade_records.append([long_pos.datetime, self.price, u'sell'])

def short_pos_cover(self, short_pos):##持有空仓时的平仓考虑。

    if short_pos.askPrice1 > (1+self.zhisun)*self.rec_price: ##做空的时候，实时价格高于止损线
        self.cover(self.askprice1, 1)
        self.pos_rec += 1
        self.writeCtaLog(u'空仓'+u'止损价'+str((1 + self.zhisun) * self.rec_price)+u'当前价'+str(self.price))
        self.zhisun_label = True
        self.zhisun_bar = 0
        self.trade_records.append([short_pos.datetime, self.price, u'cover'])

    if short_pos.askPrice1 < self.rec_price*(1-self.zhiying):##做空的时候，实时价格已经低于止盈线了
        self.cover(self.askprice1, 1)
        self.pos_rec += 1
        self.writeCtaLog(u'空仓'+u'止盈价'+str(self.rec_price * (1 - self.zhiying))+u'当前价'+str(self.price))
        self.trade_records.append([short_pos.datetime, self.price, u'cover'])

```

(6) 优化策略：防止开多仓的策略；以及亏损后考虑停止 5 分钟再次建仓，防止立刻反向操作再次亏损的情况；不留过夜仓的策略。

```

def pos_rec_concert(self): # pos与pos_rec不一致时进行调整。
    if self.pos != self.pos_rec:
        self.writeCtaLog(u'调整pos_rec, 由' + str(self.pos_rec) + u'变为' + str(self.pos))
        self.pos_rec = self.pos

    if self.lastOrder is not None and self.lastOrder.status == u'未成交':
        self.cancelOrder(self.lastOrder.vtOrderID)
        self.lastOrder = None
        self.writeCtaLog(u'撤销上一单')

```

```

#每次止损平仓后，考虑停止5分钟进行操作，因此设置zhisun_label作为是否能重新建仓的标志
def zhisun_set(self):
    if self.zhisun_label == True:
        self.zhisun_bar += 1

    if self.zhisun_bar >= 5:
        self.zhisun_bar = 0
        self.zhisun_label = False

```

```

if bar.datetime.hour == 14 and bar.datetime.minute >= 58:
    print(u'stop time')
    if self.pos == 0:
        pass
    if self.pos > 0:
        self.sell(self.bidprice1, 1)
        self.zhisun_label = True
        self.trade_records.append([bar.datetime, self.price, u'sell'])
    if self.pos < 0:
        self.cover(self.askprice1, 1)
        self.zhisun_label = True
        self.trade_records.append([bar.datetime, self.price, u'cover'])

```

从以上分析可以看出，该策略十分依赖于预测的准确性。若预测结果出现太多大涨大跌的情况，则会频繁开仓；若预测结果过于平稳，则会出现不交易的情况。若预测趋势与实际趋势相反，则会造成较大损失。

2.执行结果

(1) 实时分类结果与得分显示：

inited	trading	pos	pos_rec	count	count_bar	class_1	class_2	class_3	class_4	class_5	score
True	True	1	0	33	33	3	1	1	1	1	1.8

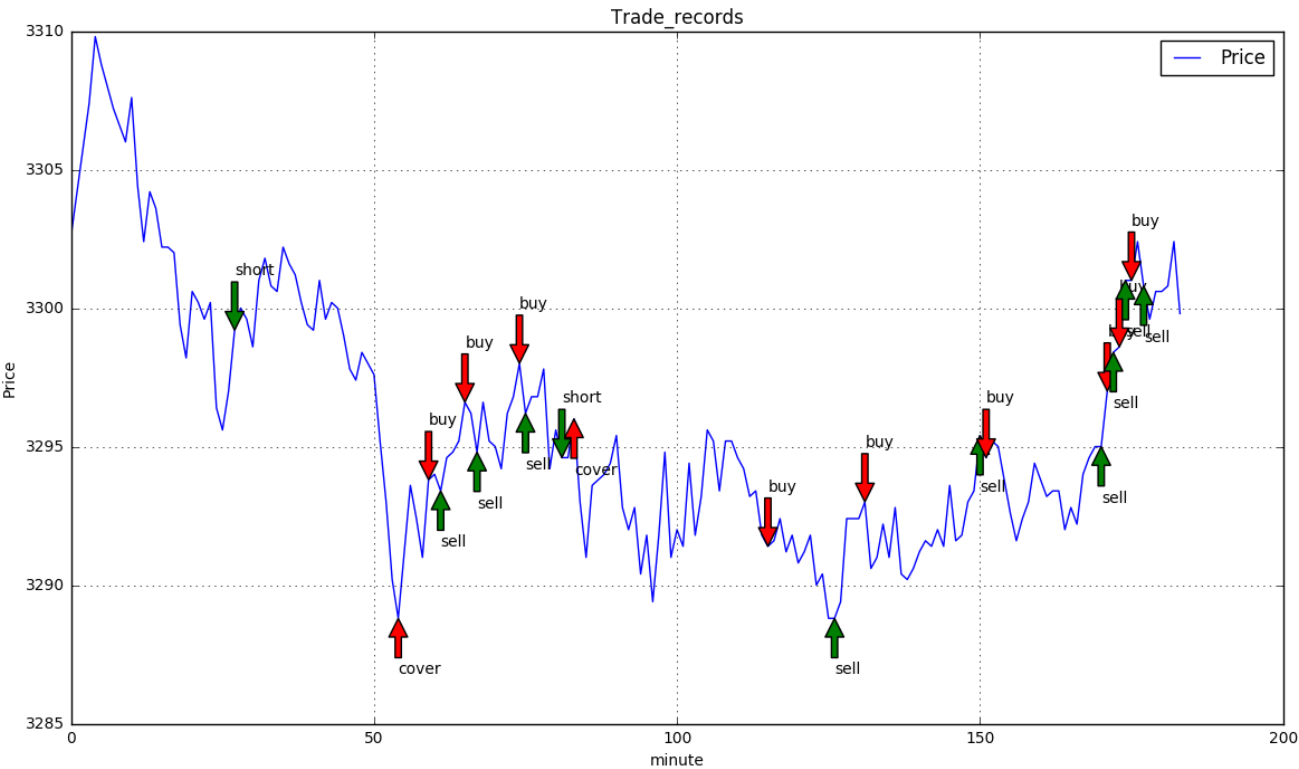
(2) 日内交易情况表（部分）

71617	8	IF1701	沪深300股指1701	空	平仓	3296.0	1	10:51:28	CTP
71086	7	IF1701	沪深300股指1701	多	开仓	3298.0	1	10:51:02	CTP
67796	6	IF1701	沪深300股指1701	空	平仓	3294.4	1	10:43:32	CTP
66981	5	IF1701	沪深300股指1701	多	开仓	3297.0	1	10:42:01	CTP
65427	4	IF1701	沪深300股指1701	空	平仓	3291.2	1	10:37:13	CTP
64967	3	IF1701	沪深300股指1701	多	开仓	3293.8	1	10:36:01	CTP
62450	2	IF1701	沪深300股指1701	多	平仓	3290.6	1	10:30:12	CTP
56610	1	IF1701	沪深300股指1701	空	开仓	3298.8	1	10:04:01	CTP

(3) 日内资金额

账户	昨结	净值	可用	手续费	保证金	平仓盈亏	持仓盈亏	接口
079240	996113.66	995611.1639	994591.1639	1522.4961		1020.0		CTP

(4) 日内交易情况图



3.模型和策略分析

目前来看，执行结果不太令人满意。主要不足有：

- (1) 预测结果有太多的大涨大跌趋势，导致频繁开仓。
- (2) 预测趋势准确率较低。从交易图示来看，11 笔交易中，趋势预测准确的只有 5 笔，未过半。
- (3) 止盈止损系数设置的过低，稍有波动就会平仓，导致持仓时间过短。如第二笔，预测是正确的，第一分钟下跌，后几分钟都是上涨。但第一分钟就止损平仓了，未能抓住大趋势。

二、策略二 ctaLSTM_V2.0（对应预测模型 LSTM_V4.0）

1.策略更改

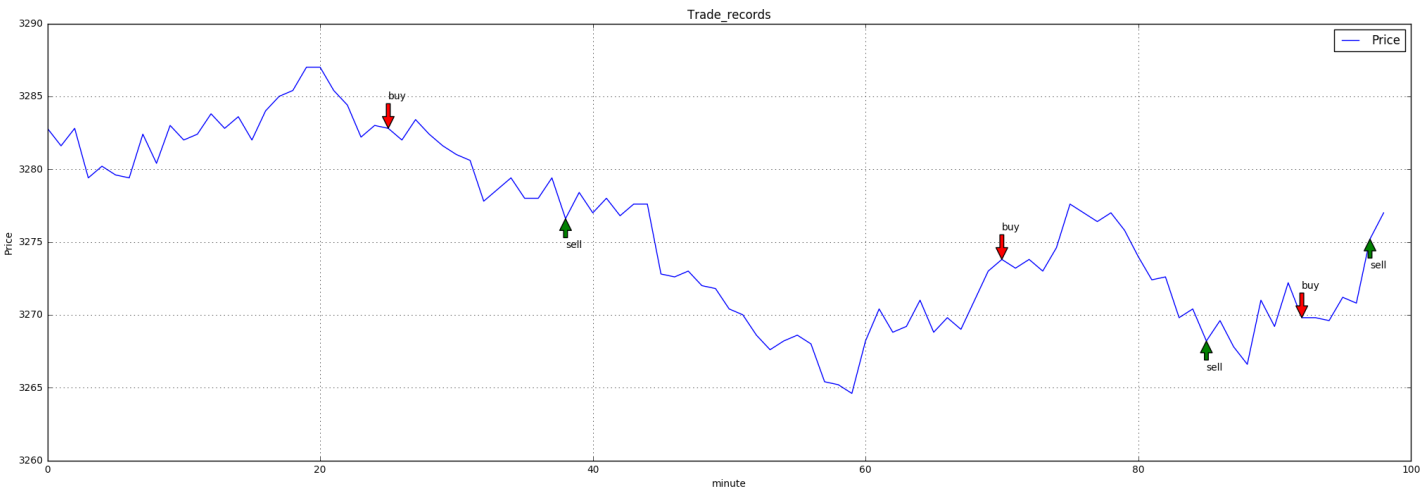
本策略的修改主要为模型的修改，从增大数据量、打乱数据、将时间序列由 5 分钟增长至 25 分钟这三个方面进行了修改。训练完之后，准确率有了极大的提高。

同时，在策略方面增加当前时刻涨跌情况和准确率显示模块，如下图：

inited	trading	pos	pos_rec	count	count_bar	real_class	predict_1	predict_2	predict_3	predict_4	predict_5	score	acc
True	True	0	0	42	42	3	3	3	5	5	5	3.8	6

其中，real_class 是当前时刻的实际分类，predict_1 至 predict_5 为预测的后 1-5 分钟的分类。acc 是今天所有分钟的 predict_1 与 real_class 的比值，即后一分钟预测成功率。

2.日内交易情况图



3.模型和策略分析

- (1) 修改后的模型在测试时性能有了明显改善，但是在实际应用中还是效果不好。之前是太多的出现大涨大跌的预测，现在是预测结果不灵敏，经常连续几分钟始终保持同样的预测结果。怀疑是 IC 和 IH 的涨跌幅与 IF 的涨跌幅区间不同，导致分类阈值设置的不准确有关。考虑更多的使用 IF 合约，加上前几年的和最近时间的。
- (2) 策略上，目前只能预测后五分钟的涨跌情况，即使预测准确率够高，开仓后要保证足够的持仓时间（大于五分钟），当出现前面几分钟涨后面几分钟大跌的情况，也是无法盈利的。另一方面，预测的时间过长，准确率肯定不会高。因此还需要对策略进行修改，确定更合理的建仓时间。

三、策略三 ctaLSTM_V3.0 (对应预测模型 LSTM_V5.0)

1.策略更改

策略三在模型方面和策略方面都进行了比较大的改动。模型上的改动有：

(1) 继续增大数据量。用上了数据源中的所有分钟数据，从 14 年到 16 年 11 月，合约包括 IC、IF、IH 三种，经过清洗后的数据量为 33 万条。

(2) 分类改变。之前是将三种合约合起来分类，由于三种合约涨跌幅差异比较大，导致分类效果很差。现在使用归一化后的结果进行分类，并且将三种合约区分开，避免合约不同带来的差异影响。同时，细化分类，由 5 类变为 7 类。统计每种合约的区间分布，使三种合约的分类比保持基本相同。如下图：

IF合约：

```
count      146399.000000
mean        0.462578
std         0.092574
min         0.000000
25%         0.432531
50%         0.461256
75%         0.489928
max         1.000000
Name: RaiseDown_1, dtype: float64
涨幅3: 1849 0.0126298676904
涨幅2: 4056 0.0277051072753
涨幅1: 18540 0.126640209291
平稳: 98495 0.672784650168
跌幅1: 17003 0.116141503699
跌幅2: 4666 0.031871802403
跌幅3: 1790 0.0122268594731
```

IC合约：

```
count      92931.000000
mean        0.443013
std         0.124230
min         0.000000
25%         0.375561
50%         0.439235
75%         0.505320
max         1.000000
Name: RaiseDown_1, dtype: float64
涨幅3: 1207 0.0129881309789
涨幅2: 2838 0.0305387868419
涨幅1: 11337 0.121993737289
平稳: 62569 0.673284479883
跌幅1: 11079 0.11921748394
跌幅2: 2870 0.0308831283425
跌幅3: 1031 0.0110942527251
```

IH合约：

```
count      95090.000000
mean        0.408681
std         0.071421
min         0.000000
25%         0.393536
50%         0.406805
75%         0.419971
max         1.000000
Name: RaiseDown_1, dtype: float64
涨幅3: 数量: 975 比例: 0.0102534441056
涨幅2: 数量: 3084 比例: 0.0324324324324
涨幅1: 数量: 11522 比例: 0.121169418446
平稳: 数量: 63855 比例: 0.671521716269
跌幅1: 数量: 11419 比例: 0.120086234094
跌幅2: 数量: 3219 比例: 0.0338521400778
跌幅3: 数量: 1016 比例: 0.0106846145757
```

(3) 更改涨跌幅计算方法。如计算第四分钟涨跌幅，之前是计算第四分钟相对于第一分钟的涨跌，现在是计算第四分钟相对于第三分钟的涨跌。这样能够更好的分析交易情况图，并且大大减少大涨大跌情况出现的频率。

(4) 更改训练集和测试集的比重。由 8:2 改为 7:3。后者是公认最好的训练测试集比重。

策略上的改动为：

(1) 修改平仓条件。原来是完全根据止盈止损策略进行平仓，现在完全根据预测结果进行平仓，以体现预测的重要性。（后面陆续添加止损策略和止盈策略）

(2) 增加了一个开仓条件。原来是完全根据预测得分开仓，现在加了后一分钟的涨跌分类限制，使开仓条件更为严苛，防止盲目开仓。

(3) 更改保存数据的代码段位置。之前是点击“停止策略”后保存当天数据，现在是每接收到一个 bar 数据自动保存，避免由于死机造成数据无法保存的情况。

(4) 根据预测模型中 7 种分类的分类比，修改当天实际数据的分类阈值。

2. 执行结果

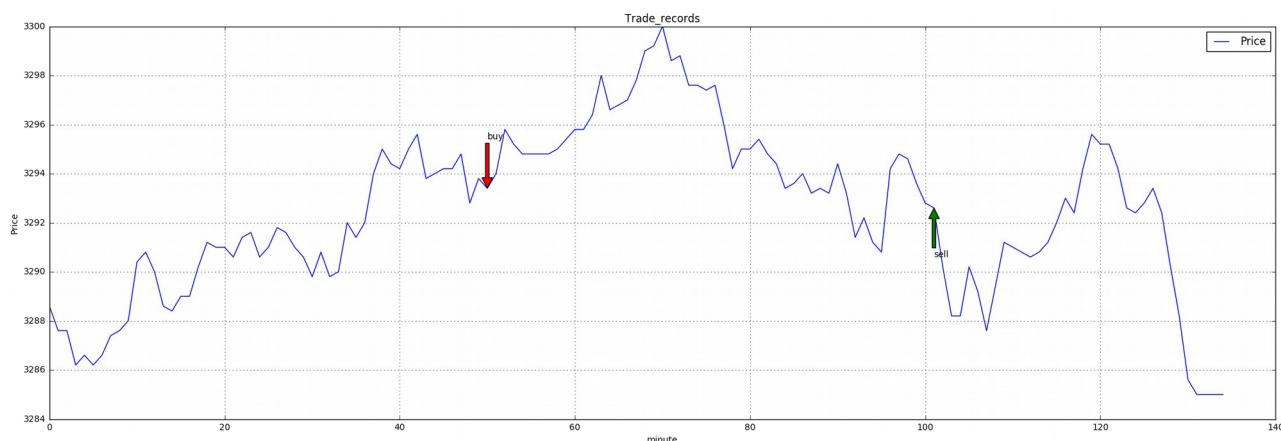
(1) 分类动态显示

inited	trading	pos	pos_rec	count	count_bar	real_class	predict_1	predict_2	predict_3	predict_4	predict_5	score	acc
True	True	1	1	97	97	7	5	7	3	2	1	4.3	0.527777...

从一下午的情况来看，预测的分类不再出现很多大涨大跌的情况，也不会一直保持平稳预测，更加贴近实际情况了。

但是，准确率还是远远不足。一分钟预测的准确率只能达到 50% 多一些，其中平稳的预测占了大多数，难以实现稳定的盈利。

(2) 日内交易图



可以看出，做出了两次预测，都是比较正确的。但由于在最高点处没有准确预测出下跌，导致预测下跌时已经低至了一定价格。

3. 模型和策略分析

预测准确率还不能使人满意，需要不断探索更好的模型结构，使准确率不断提高。

另外，为了更好的反映出预测是否有效，正在尝试添加策略代码，使其能够将预测数据输出，进行进一步的分析，并有针对性的改进。