Research in Collective Intelligence through Horse Racing in Hong Kong

LYU1604
Lam Cheuk Hei (1155049403)
Chan Lok Wai (1155050024)
Supervised By Prof. Michael R. Lyu
Outline

Introduction
Implementation
Research Results
Discussion and Recommendations
Demonstration
Conclusion
Introduction - Term 1 Recap
Introduction - Term 1 Recap

Problems:

- random predictions
- may not provide tips regularly
- algorithm may not work

Solution:

- change target audience - horse racing experts
Introduction - Term 2 Design

Collecting data

Website

Tips on newspaper
Introduction - Term 2 Design

Analyze Data

TensorFlow

Voting System
What is collective intelligence

Collect and combine information from two or more people

Provide solution to a specific problem
What is machine learning

gives computers the ability to learn without being explicitly programmed
related to computational statistics
focuses on prediction-making through the use of computers
What is TensorFlow

machine learning algorithms

open source software

implementation for executing such algorithms
Win and Place?

basic bet types for horse racing

Win

win if the chosen horse finish first

average winning chance

Place

win if the chosen horse finish first, second or third

high winning chance
Research Participants

54 experts who provide tips on newspaper
## Research Participants

<table>
<thead>
<tr>
<th>日期</th>
<th>星期</th>
<th>第一次</th>
<th>第二次</th>
<th>第三次</th>
<th>第四次</th>
</tr>
</thead>
<tbody>
<tr>
<td>2017-01-25</td>
<td>三</td>
<td>1</td>
<td>6</td>
<td>11</td>
<td>4</td>
</tr>
<tr>
<td>2017-01-26</td>
<td>三</td>
<td>2</td>
<td>4</td>
<td>9</td>
<td>2</td>
</tr>
<tr>
<td>2017-01-27</td>
<td>三</td>
<td>3</td>
<td>7</td>
<td>5</td>
<td>9</td>
</tr>
<tr>
<td>2017-01-28</td>
<td>三</td>
<td>4</td>
<td>1</td>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td>2017-01-29</td>
<td>三</td>
<td>5</td>
<td>9</td>
<td>10</td>
<td>8</td>
</tr>
<tr>
<td>2017-01-30</td>
<td>三</td>
<td>6</td>
<td>9</td>
<td>8</td>
<td>10</td>
</tr>
<tr>
<td>2017-01-31</td>
<td>三</td>
<td>7</td>
<td>6</td>
<td>3</td>
<td>7</td>
</tr>
<tr>
<td>2017-02-01</td>
<td>三</td>
<td>8</td>
<td>12</td>
<td>4</td>
<td>7</td>
</tr>
<tr>
<td>2017-02-02</td>
<td>三</td>
<td>9</td>
<td>9</td>
<td>8</td>
<td>4</td>
</tr>
<tr>
<td>2017-02-03</td>
<td>三</td>
<td>1</td>
<td>6</td>
<td>9</td>
<td>8</td>
</tr>
<tr>
<td>2017-02-04</td>
<td>三</td>
<td>2</td>
<td>1</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>2017-02-05</td>
<td>三</td>
<td>3</td>
<td>9</td>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td>2017-02-06</td>
<td>三</td>
<td>4</td>
<td>7</td>
<td>3</td>
<td>10</td>
</tr>
<tr>
<td>2017-02-07</td>
<td>三</td>
<td>5</td>
<td>4</td>
<td>10</td>
<td>1</td>
</tr>
<tr>
<td>2017-02-08</td>
<td>三</td>
<td>6</td>
<td>14</td>
<td>6</td>
<td>13</td>
</tr>
<tr>
<td>2017-02-09</td>
<td>三</td>
<td>7</td>
<td>14</td>
<td>9</td>
<td>6</td>
</tr>
<tr>
<td>2017-02-10</td>
<td>三</td>
<td>8</td>
<td>3</td>
<td>7</td>
<td>4</td>
</tr>
<tr>
<td>2017-02-11</td>
<td>三</td>
<td>9</td>
<td>1</td>
<td>13</td>
<td>2</td>
</tr>
<tr>
<td>2017-02-12</td>
<td>三</td>
<td>10</td>
<td>12</td>
<td>10</td>
<td>13</td>
</tr>
<tr>
<td>2017-02-13</td>
<td>三</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>12</td>
</tr>
<tr>
<td>2017-02-14</td>
<td>三</td>
<td>2</td>
<td>7</td>
<td>11</td>
<td>3</td>
</tr>
<tr>
<td>2017-02-15</td>
<td>三</td>
<td>3</td>
<td>3</td>
<td>5</td>
<td>7</td>
</tr>
<tr>
<td>2017-02-16</td>
<td>三</td>
<td>4</td>
<td>10</td>
<td>8</td>
<td>4</td>
</tr>
<tr>
<td>2017-02-17</td>
<td>三</td>
<td>5</td>
<td>5</td>
<td>6</td>
<td>9</td>
</tr>
<tr>
<td>2017-02-18</td>
<td>三</td>
<td>6</td>
<td>12</td>
<td>10</td>
<td>8</td>
</tr>
<tr>
<td>2017-02-19</td>
<td>三</td>
<td>7</td>
<td>1</td>
<td>5</td>
<td>14</td>
</tr>
<tr>
<td>2017-02-20</td>
<td>三</td>
<td>8</td>
<td>1</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>
## Research Participants

<table>
<thead>
<tr>
<th>Date</th>
<th>No.1</th>
<th>No.2</th>
<th>No.3</th>
<th>No.4</th>
</tr>
</thead>
<tbody>
<tr>
<td>2017-01-25</td>
<td>1</td>
<td>6</td>
<td>11</td>
<td>4</td>
</tr>
<tr>
<td>2017-01-26</td>
<td>2</td>
<td>4</td>
<td>9</td>
<td>2</td>
</tr>
<tr>
<td>2017-01-27</td>
<td>3</td>
<td>7</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>2017-01-28</td>
<td>4</td>
<td>1</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>2017-01-29</td>
<td>5</td>
<td>9</td>
<td>10</td>
<td>8</td>
</tr>
</tbody>
</table>

---

### Results

1. Red for first place, blue for second place, green for third place, yellow for fourth place.
2. If the odd number of valid participants appears, it indicates their presence at the event.

###Various Reports

- **Eastern Daily**: Red for first place, blue for second place, green for third place, yellow for fourth place.
- **Taiwan Times**: Red for first place, blue for second place, green for third place, yellow for fourth place.
- **Morning Star**: Red for first place, blue for second place, green for third place, yellow for fourth place.
- **Asia Daily**: Red for first place, blue for second place, green for third place, yellow for fourth place.
- **Central Times**: Red for first place, blue for second place, green for third place, yellow for fourth place.
- **Eagles Times**: Red for first place, blue for second place, green for third place, yellow for fourth place.

---

### Various Reports

- **Eastern Daily**: 10:236, 8:237, 7:238, 5:239
- **Taiwan Times**: 7:238, 5:239, 3:240, 1:241
- **Morning Star**: 2:242, 1:243, 4:244, 5:245
- **Asia Daily**: 6:246, 5:247, 3:248, 1:249
- **Central Times**: 8:250, 6:251, 4:252, 2:253
- **Eagles Times**: 2:254, 1:255, 3:256, 5:257

---

### Additional Information

- **Eastern Daily**: 425, 8236, 7237, 5238, 3239
- **Taiwan Times**: 7238, 5239, 3240, 1241
- **Morning Star**: 2242, 1243, 4244, 5245
- **Asia Daily**: 6246, 5247, 3248, 1249
- **Central Times**: 8250, 6251, 4252, 2253
- **Eagles Times**: 2254, 1255, 3256, 5257

---

### Contact Information

For inquiries, please contact: info@hohowen.com or call (852) 3168 6978 (886) 1471 102 8888 or visit www.hohowen.com.
Voting System

All experts vote for horses

Rank the horses by votes

Special case:

horses with same votes

Solution:

we choose the one with lower horse number

Reason:

Lower horse number win: 89%

Higher horse number win: 11%
Transform tips to horse racing database
add 1 if the expert predicted that horse
new column "result":
1 if the place\leq3
-1 if the place>3
TensorFlow - features

distance
course
going
jockey code
trainer code
actual weight
horse weight
4 experts’ tips
result
TensorFlow - number of training and predicting races

All tips: 569 races
Training: 452 races
Predicting: 117 races
TensorFlow - configuration

128 neural networks
## TensorFlow - prediction

<table>
<thead>
<tr>
<th>no</th>
<th>jockeycode</th>
<th>trainercode</th>
<th>actualweight</th>
<th>horseweight</th>
<th>winodds</th>
<th>od28</th>
<th>ad42</th>
<th>st47</th>
<th>sp51</th>
<th>result</th>
<th>prediction</th>
<th>predict_place</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>LHW</td>
<td>MA</td>
<td>118</td>
<td>1048</td>
<td>36</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-1.85967</td>
<td>12</td>
</tr>
<tr>
<td>1</td>
<td>CSF</td>
<td>JM</td>
<td>133</td>
<td>1128</td>
<td>13</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-1</td>
<td>-0.772455</td>
<td>10</td>
</tr>
<tr>
<td>2</td>
<td>CKK</td>
<td>YCH</td>
<td>127</td>
<td>1106</td>
<td>30</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>-1</td>
<td>0.596704</td>
<td>6</td>
</tr>
<tr>
<td>3</td>
<td>DO</td>
<td>CCW</td>
<td>119</td>
<td>1067</td>
<td>12</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-1</td>
<td>0.652511</td>
<td>8</td>
</tr>
<tr>
<td>4</td>
<td>LDE</td>
<td>YPF</td>
<td>114</td>
<td>1097</td>
<td>99</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-1</td>
<td>-0.608304</td>
<td>7</td>
</tr>
<tr>
<td>5</td>
<td>NGA</td>
<td>CAS</td>
<td>111</td>
<td>1142</td>
<td>40</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-1</td>
<td>0.672555</td>
<td>9</td>
</tr>
<tr>
<td>6</td>
<td>STH</td>
<td>FD</td>
<td>128</td>
<td>1039</td>
<td>99</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-1</td>
<td>0.782188</td>
<td>11</td>
</tr>
<tr>
<td>7</td>
<td>PZ</td>
<td>FC</td>
<td>128</td>
<td>1097</td>
<td>38</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-1</td>
<td>-0.490287</td>
<td>5</td>
</tr>
</tbody>
</table>
Combined Method

- Voting system
- Combined method
- TensorFlow
Accuracy

ad42
od28
sp51
st47
# Accuracy - Win

117 races

117 horses

<table>
<thead>
<tr>
<th>Tips provider</th>
<th>Number of correct prediction</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>ad42</td>
<td>30</td>
<td>25.6%</td>
</tr>
<tr>
<td>od28</td>
<td>33</td>
<td>28.2%</td>
</tr>
<tr>
<td>sp51</td>
<td>30</td>
<td>25.6%</td>
</tr>
<tr>
<td>st47</td>
<td>20</td>
<td>17.1%</td>
</tr>
<tr>
<td>voting system</td>
<td>48</td>
<td>41.0%</td>
</tr>
<tr>
<td>TensorFlow</td>
<td>38</td>
<td>32.5%</td>
</tr>
<tr>
<td>combined method</td>
<td>26 (in 57 races)</td>
<td>45.6% (in 57 races)</td>
</tr>
</tbody>
</table>
Accuracy - Win

Accuracy for win

- ad42
- od28
- sp51
- st47
- voting system
- tensorflow
- combined method

Accuracy
## Accuracy - Place

117 races

351 horses

<table>
<thead>
<tr>
<th>Tips provider</th>
<th>Number of correct prediction</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>ad42</td>
<td>166</td>
<td>47.3%</td>
</tr>
<tr>
<td>od28</td>
<td>169</td>
<td>48.1%</td>
</tr>
<tr>
<td>sp51</td>
<td>154</td>
<td>43.9%</td>
</tr>
<tr>
<td>st47</td>
<td>165</td>
<td>47.0%</td>
</tr>
<tr>
<td>voting system</td>
<td>192</td>
<td>54.7%</td>
</tr>
<tr>
<td>TensorFlow</td>
<td>172</td>
<td>49.0%</td>
</tr>
<tr>
<td>combined method</td>
<td>138 (out of 235 horses)</td>
<td>58.7%      (out of 235 horses)</td>
</tr>
</tbody>
</table>
Accuracy - Place

Accuracy for place

- ad42
- od28
- sp51
- st47
- voting system
- tensorflow
- combined method

Accuracy
Gain and loss analysis

Experts, voting system, TensorFlow and combined method

Total 117 races
Gain and loss analysis - 1 win

Choose one horse in each race, and buy ‘win’ for it

$10 for each horse

Total cost is $1170 (Except Combined method)

Combined method: only bet on 57 races over 117 races, total cost is $570
Gain and loss analysis - 3 place

Choose three horses in each race, and buy ‘place’ for all of them

$10 for each horse

$30 for each race

Total cost is $3510 (Except Combined method)

Combined method:

only bet on 235 horses in 114 races over 117 races, total cost is $2350
Gain and loss analysis

Voting system $176.2
Gain and loss analysis - 1 win without TensorFlow

Choose one horse in each race, and buy ‘win’ for it

Total 569 races

$10 for each horse

Total cost is $5690
Cumulative net gain and loss analysis of '1 win' over 569 races

Voting system $1086
Profit and loss percentage

Combined method can’t get highest net gain
But it doesn’t mean that combined method is not good
Profit and loss percentage=net gain over money spent
Profit and loss percentage - 1 win

Combined method 55.3%
Profit and loss percentage - 3 place

Combined method 5.02%
Maximization of returns

Combined method (highest profit percentage)

1 win

Further improvement:

Only bet on horses with win odds lower than 10
Maximization of returns - Result

1 win

bet on 95 races over 117 races

For combined method: only bet on 48 races
Cumulative net gain of '1 win' (win odds < 10)

voting system $766
Combined method 84.4%
## Maximization of returns - Result

<table>
<thead>
<tr>
<th>Tips provider</th>
<th>Net gain</th>
<th>profit and loss percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Before</td>
<td>After</td>
</tr>
<tr>
<td>ad42</td>
<td>102</td>
<td>112</td>
</tr>
<tr>
<td>od28</td>
<td>-167</td>
<td>53</td>
</tr>
<tr>
<td>sp51</td>
<td>-147</td>
<td>73</td>
</tr>
<tr>
<td>st47</td>
<td>-495</td>
<td>-275</td>
</tr>
<tr>
<td>voting system</td>
<td>546</td>
<td>766</td>
</tr>
<tr>
<td>TensorFlow</td>
<td>410</td>
<td>510</td>
</tr>
<tr>
<td>Combined method</td>
<td>315</td>
<td>405 (in 48 races)</td>
</tr>
</tbody>
</table>
Let’s see our prediction on 20-4-2017

<table>
<thead>
<tr>
<th>race no</th>
<th>1st</th>
<th>2nd</th>
<th>3rd</th>
<th>predicted winner from voting system</th>
<th>Win Odds</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5</td>
<td>2</td>
<td>10</td>
<td>5</td>
<td>2.5</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>2</td>
<td>8</td>
<td>3</td>
<td>5.7</td>
</tr>
<tr>
<td>3</td>
<td>7</td>
<td>9</td>
<td>4</td>
<td>9</td>
<td>4.7</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>7</td>
<td>12</td>
<td>1</td>
<td>4.6</td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td>10</td>
<td>5</td>
<td>2</td>
<td>4.6</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>8</td>
<td>2</td>
<td>1</td>
<td>4.7</td>
</tr>
<tr>
<td>7</td>
<td>8</td>
<td>4</td>
<td>2</td>
<td>8</td>
<td>5.6</td>
</tr>
<tr>
<td>8</td>
<td>3</td>
<td>12</td>
<td>9</td>
<td>7</td>
<td>16</td>
</tr>
</tbody>
</table>

We don’t bet on race 8 because the win odd was higher than 10 before the race.

We bet $10 on each horse.

We predicted 5 races correctly over 7 races.

Net gain=$147 in one single day.
Conclusion

Is crowdsourcing better than individual experts?

Yes

What is the best strategy to bet on horse racing?

‘1 win’ with win odds lower than 10 using combined method
increase the betting amount if you want to earn more money
Thank You