
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

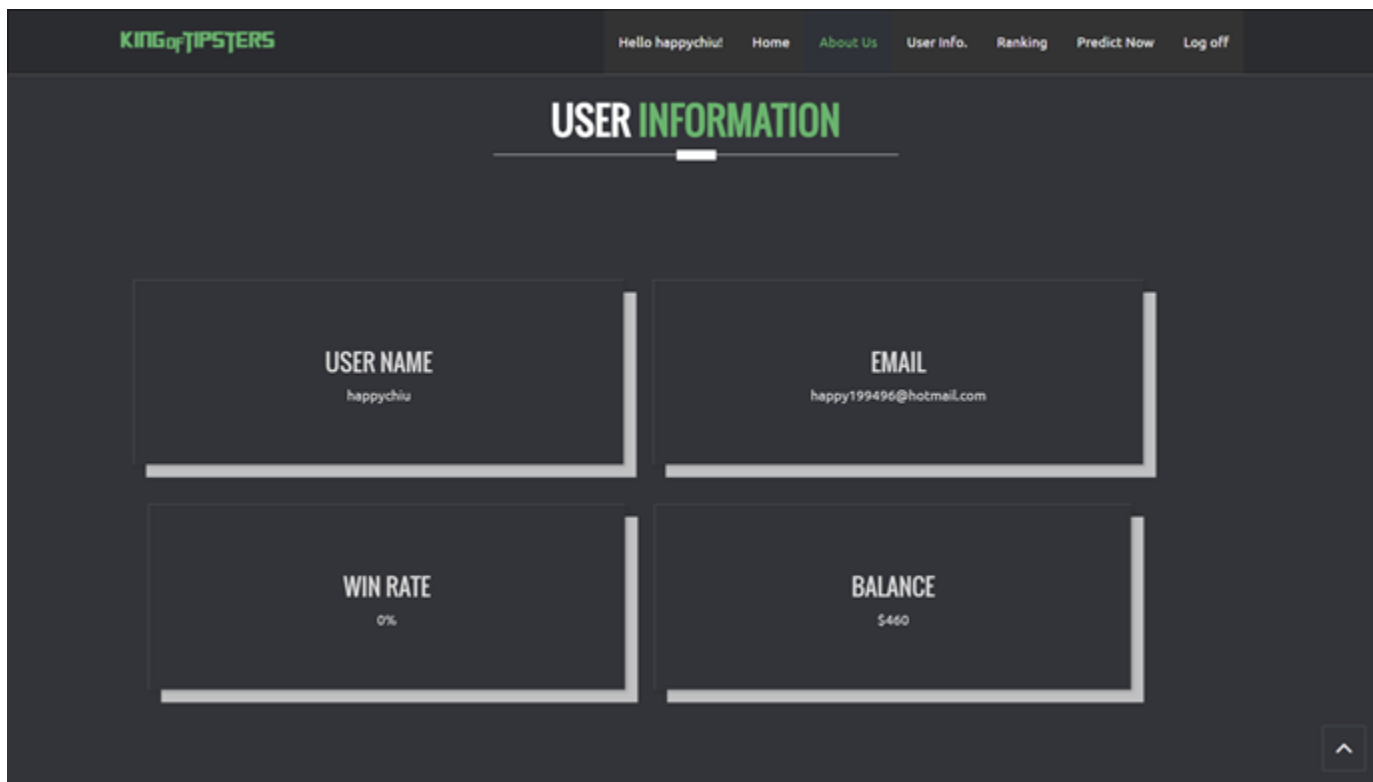
Research Results

Discussion and Recommendations

Demonstration

Conclusion

Introduction - Term 1 Recap



The screenshot shows a dark-themed user profile page. At the top left is the logo 'KING of TIPSTERS'. The top navigation bar includes 'Hello happychiu', 'Home', 'About Us', 'User Info.', 'Ranking', 'Predict Now', and 'Log off'. The main heading is 'USER INFORMATION' with a horizontal line underneath. Below this are four data cards: 'USER NAME' (happychiu), 'EMAIL' (happy199496@hotmail.com), 'WIN RATE' (0%), and 'BALANCE' (\$460). A small upward arrow icon is in the bottom right corner.

KING of TIPSTERS

Hello happychiu | Home | About Us | User Info. | Ranking | Predict Now | Log off

USER INFORMATION

USER NAME happychiu	EMAIL happy199496@hotmail.com
WIN RATE 0%	BALANCE \$460

↑

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

日期	星期	場次	No.1	No.2	No.3	No.4
2017-01-25 40期	三	1	6	11	4	9
2017-01-25 40期	三	2	4	9	2	8
2017-01-25 40期	三	3	7	5	9	2
2017-01-25 40期	三	4	1	2	7	3
2017-01-25 40期	三	5	9	10	8	12
2017-01-25 40期	三	6	9	8	10	2
2017-01-25 40期	三	7	6	3	7	1
2017-01-25 40期	三	8	12	4	7	8
2017-01-25 40期	三	9	9	8	4	6
2017-01-22 39期	日	1	6	3	8	1
2017-01-22 39期	日	2	1	4	3	9
2017-01-22 39期	日	3	9	5	10	6
2017-01-22 39期	日	4	7	3	10	4
2017-01-22 39期	日	5	4	10	1	13
2017-01-22 39期	日	6	14	6	13	4
2017-01-22 39期	日	7	14	9	6	13
2017-01-22 39期	日	8	3	7	4	1
2017-01-22 39期	日	9	1	13	2	9
2017-01-22 39期	日	10	12	10	13	9
2017-01-18 38期	三	1	2	5	12	9
2017-01-18 38期	三	2	7	11	3	10
2017-01-18 38期	三	3	5	7	1	2
2017-01-18 38期	三	4	10	8	4	9
2017-01-18 38期	三	5	5	6	9	3
2017-01-18 38期	三	6	7	2	3	9
2017-01-18 38期	三	7	1	5	14	11
2017-01-18 38期	三	8	1	3	4	5

X	東方日報/**	派彩	X	蘋果日報**	派彩	X	明報/**	派彩	X	新報領先/**	派彩	X	成報/**	派彩					
4	2	5	6	236	4	5	2	7	4	7	5	9	4	5	1	12			
2	4	8	1	55	2	4	7	9	55	2	6	1	5	2	4	9	6	5	
3	8	4	7	35.5	7	4	2	5	35.5	7	6	5	4	35.5	7	2	6	5	3
4	2	7	8		2	4	8	6		6	5	2	4		2	4	8	10	
9	3	1	8	22	9	1	3	10	22	9	1	3	11	22	3	9	11	1	22
10	1	7	6		11	1	7	8		1	4	10	2		1	9	10	6	43.5
1	3	7	9		6	9	1	3	36	3	2	1	7		2	6	1	3	36
7	8	11	5		8	2	11	7		8	5	2	7		8	2	11	1	
2	8	9	4	26	4	9	8	1	26	2	8	4	9	26	9	8	2	3	26
1	6	8	10	53.5	13	8	6	10	53.5	6	1	8	9	53.5	13	6	8	10	53.5
1	12	4	7	23.5	1	9	6	13	23.5	3	1	6	2	23.5	1	2	5	13	23.5
1	6	2	3		6	3	11	5		6	3	1	2		6	3	2	1	
7	5	4	3	31	1	3	4	10		3	7	2	1	31	3	4	9	1	
11	2	3	10		3	2	10	11		11	2	6	10		3	2	10	14	
6	14	10	12	64.5	12	14	13	6	64.5	10	14	6	13	64.5	14	6	13	2	64.5
14	7	1	2	17	4	14	1	8	17	14	6	2	7	17	1	3	14	9	17
7	11	10	8		11	7	1	10		7	11	10	2		7	2	1	11	
11	1	5	2	21	3	11	1	2	21	1	2	3	5	21	1	13	5	2	21
5	1	2	3		2	1	5	6		5	1	2	3		5	2	6	1	
2	1	4	6	28	4	2	10	3	28	4	2	3	10	28	2	4	6	10	28
6	3	5	1		6	1	4	3		6	4	3	1		6	3	2	4	
4	6	5	2	65	4	6	1	7		2	4	5	6	65	6	2	1	5	65
8	5	12	10	120	8	2	9	3		2	12	10	8	120	8	10	5	3	120
9	11	5	12	84.5	1	9	11	10		9	5	10	11	84.5	11	9	4	5	84.5
6	8	2	5		8	7	6	10	74.5	6	8	10	7	74.5	7	10	8	6	74.5
2	1	7	4	64	7	11	2	1	64	1	11	5	2	64	1	2	9	7	64
1	3	5	6	26	1	3	5	8	26	6	3	1	2	26	1	3	5	11	26

* 聰明的投注方法(不用犯本, 30倍才投注)
你估邊份報紙的專家最利害!

Research Participants

<input checked="" type="checkbox"/> 東方日報/**	64.9%	<input checked="" type="checkbox"/> 蘋果日報/**	63.5%	<input checked="" type="checkbox"/> 明報/**	63.5%	<input checked="" type="checkbox"/> 節節領先/**	63.2%	<input type="checkbox"/> 成報/**	62.2%	<input checked="" type="checkbox"/> 明報/**	61.9%	<input checked="" type="checkbox"/> 節節領先/**	61.4%
<input checked="" type="checkbox"/> 賽馬消息/**	61.1%	<input checked="" type="checkbox"/> 賽馬消息/**	61.1%	<input checked="" type="checkbox"/> 東方日報/**	61.1%	<input checked="" type="checkbox"/> 節節領先/**	60.8%	<input checked="" type="checkbox"/> 星島日報/**	60.8%	<input checked="" type="checkbox"/> 成報/**	60.8%	<input checked="" type="checkbox"/> 新報/**	60.3%
<input checked="" type="checkbox"/> 成報/**	60.3%	<input checked="" type="checkbox"/> 東方日報/**	60%	<input checked="" type="checkbox"/> 星島日報/**	60%	<input checked="" type="checkbox"/> 專業馬訊/**	59.7%	<input checked="" type="checkbox"/> 東方日報/**	59.2%	<input checked="" type="checkbox"/> 專業馬訊/**	58.9%	<input checked="" type="checkbox"/> 星島日報/**	58.9%
<input checked="" type="checkbox"/> 成報/**	58.4%	<input checked="" type="checkbox"/> 賽馬消息/**	58.1%	<input checked="" type="checkbox"/> 賽馬消息/**	58.1%	<input checked="" type="checkbox"/> 專業馬訊/**	58.1%	<input checked="" type="checkbox"/> 專業馬訊/**	57.8%	<input checked="" type="checkbox"/> 東方日報/**	57.8%	<input checked="" type="checkbox"/> 蘋果日報/**	57.6%

全選 返選 + 展開全部專家

可勾選以上1至5位貼士專家同時查詢各專家資訊(再按近日期) 如客戶想在本網刊登廣告可至電: (852)3168 6978 (86)147 1 02 8487 或電郵查詢 聯絡我們: info@hohowen.com

賽馬結果

- 註:1. 紅色為第一名, 藍色為第二名, 綠色為第三名, 黃色為第四名。
2. 如排位出現 3/4, 即表示該名次出現平頭馬。

各大報章/馬報貼士供參考及可贏派彩金額

- 註:1. 紅色為第一名, 藍色為第二名, 綠色為第三名, 黃色為第四名。
2. 派彩為: 你所觀看之彩池(派彩)

日期	星期	場次	No.1	No.2	No.3	No.4
2017-01-25 40期	三	1	6	11	4	9
2017-01-25 40期	三	2	4	9	2	8
2017-01-25 40期	三	3	7	5	9	2
2017-01-25 40期	三	4	1	2	7	3
2017-01-25 40期	三	5	9	10	8	12
2017-01-25 40期	三	6	9	8	10	2
2017-01-25 40期	三	7	6	3	7	1
2017-01-25 40期	三	8	12	4	7	8
2017-01-25 40期	三	9	9	8	4	6
2017-01-22 39期	日	1	6	3	8	1
2017-01-22 39期	日	2	1	4	3	9
2017-01-22 39期	日	3	9	5	10	6
2017-01-22 39期	日	4	7	3	10	4
2017-01-22 39期	日	5	4	4	1	13
2017-01-22 39期	日	6	14	6	13	4

X	東方日報/**	派彩	X	蘋果日報/**	派彩	X	明報/**	派彩	X	節節領先/**	派彩
4	2 5 6	236	4	5 2 7		4	1 10 7		4	7 5 9	
2	4 8 1	55	2	4 7 9	55	2	6 1 5		5	4 2 9	55
3	8 4 7	35.5	7	4 2 5	35.5	7	6 5 4	35.5	7	2 4 8	35.5
4	2 7 8		2	4 8 6		6	5 2 4		2	4 8 10	
9	3 1 8	22	9	1 3 10	22	9	1 3 11	22	3	9 11 1	22
10	1 7 6		11	1 7 8		1	4 10 2		1	9 10 6	43.5
1	3 7 9		6	9 1 3	36	3	2 1 7		2	6 1 3	36
7	8 11 5		8	2 11 7		8	5 2 7		8	2 11 1	
2	8 9 4	26	4	9 8 1	26	2	8 4 9	26	9	8 2 3	26
1	6 8 10	53.5	13	8 6 10	53.5	6	1 8 9	53.5	13	6 8 10	53.5
1	12 4 7	23.5	1	9 6 13	23.5	3	1 6 2	23.5	1	2 5 13	23.5
1	6 2 3		6	3 11 5		6	3 1 2		6	3 2 1	
7	5 4 3	31	1	3 4 10		3	7 2 1	31	3	4 9 1	
11	2 3 10		3	2 10 11		11	2 6 10		3	2 10 14	
6	14 10 12	64.5	12	14 13 6	64.5	10	14 6 13	64.5	14	6 13 2	64.5

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%

TensorFlow

Transform tips to horse racing database

add 1 if the expert predicted that horse

new column "result" :

1 if the $\text{place} \leq 3$

-1 if the $\text{place} > 3$

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

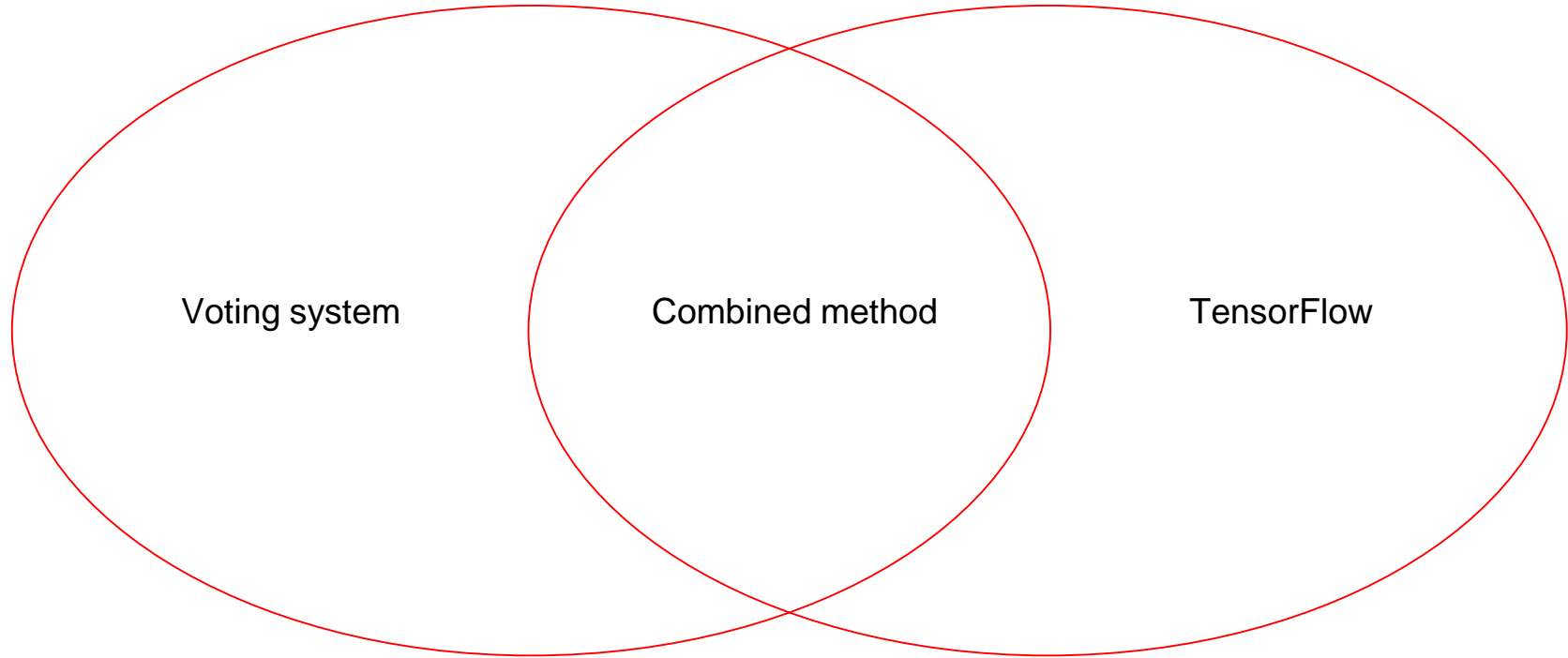
TensorFlow - configuration

128 neural networks

TensorFlow - prediction

	jockeycode	trainercode	actualweight	horseweight	winodds	od28	ad42	st47	sp51	result	prediction	predict_place
3	MHT	TYS	112	1179	2.2	1	1	1	1	1	0.150422	1
6	MOJ	SJJ	124	1116	3.8	1	1	0	0	1	-0.013719	2
5	CAN	SCS	125	1154	8.1	0	1	1	1	1	-0.034013	3
7	TEK	MKL	122	1057	10	1	0	0	1	-1	-0.284753	4
0	LHW	MA	118	1048	36	0	0	0	0	-1	-0.85967	12
1	CSF	JM	133	1128	13	0	0	0	0	-1	-0.772455	10
2	CKK	YCH	127	1106	30	0	0	1	0	-1	-0.596704	6
9	DO	CCW	119	1067	12	0	0	0	0	-1	-0.652511	8
1	LDE	YPF	114	1097	99	0	0	0	0	-1	-0.608304	7
2	NGA	CAS	111	1142	40	0	0	0	0	-1	-0.672555	9
3	STH	FD	128	1039	99	0	0	0	0	-1	-0.782188	11
4	PZ	FC	128	1097	38	0	0	0	0	-1	-0.490287	5

Combined Method



Accuracy

ad42

od28

sp51

st47

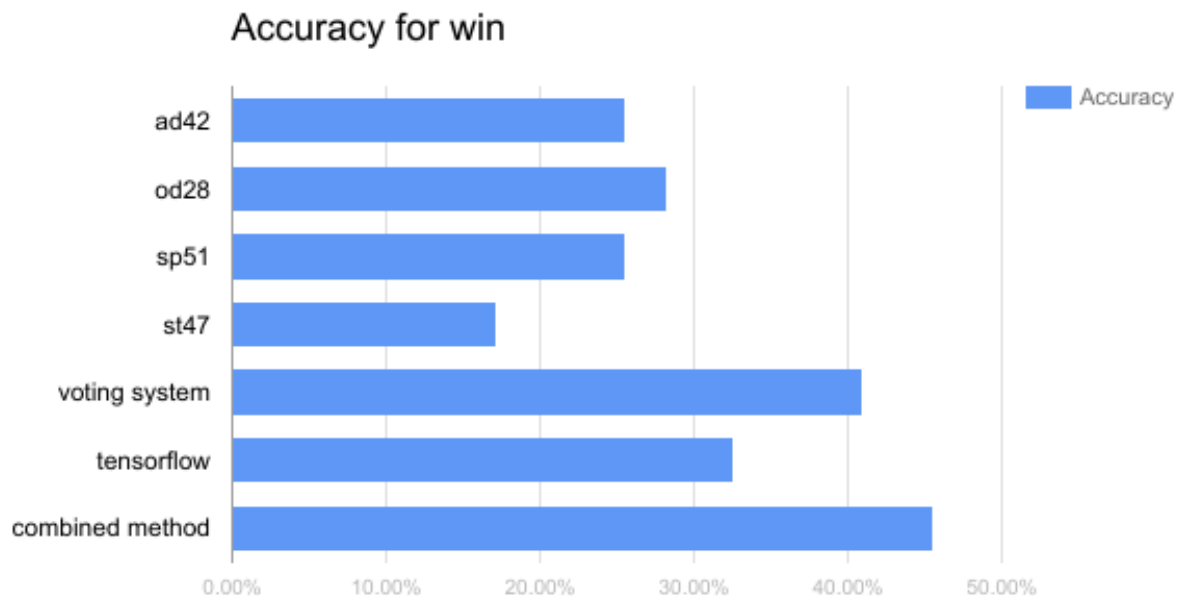
Accuracy - Win

117 races

117 horses

Tips provider	Number of correct prediction	Accuracy
ad42	30	25.6%
od28	33	28.2%
sp51	30	25.6%
st47	20	17.1%
voting system	48	41.0%
TensorFlow	38	32.5%
combined method	26 (in 57 races)	45.6% (in 57 races)

Accuracy - Win



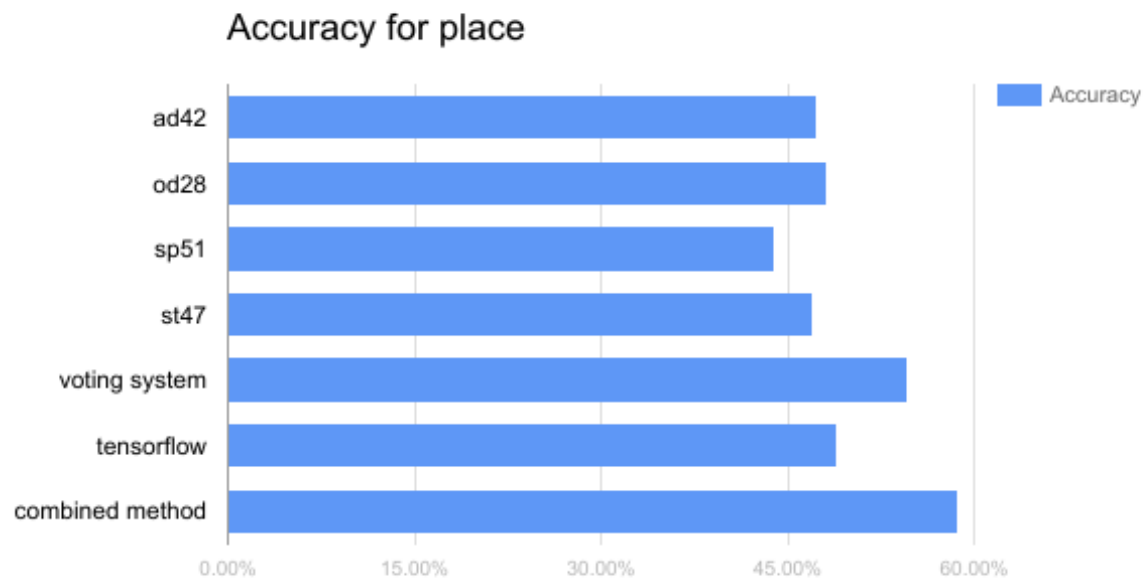
Accuracy - Place

117 races

351 horses

Tips provider	Number of correct prediction	Accuracy
ad42	166	47.3%
od28	169	48.1%
sp51	154	43.9%
st47	165	47.0%
voting system	192	54.7%
TensorFlow	172	49.0%
combined method	138 (out of 235 horses)	58.7% (out of 235 horses)

Accuracy - Place



Gain and loss analysis

Experts, voting system, TensorFlow and combined method

Total 117 races

Gain and loss analysis - 1 win

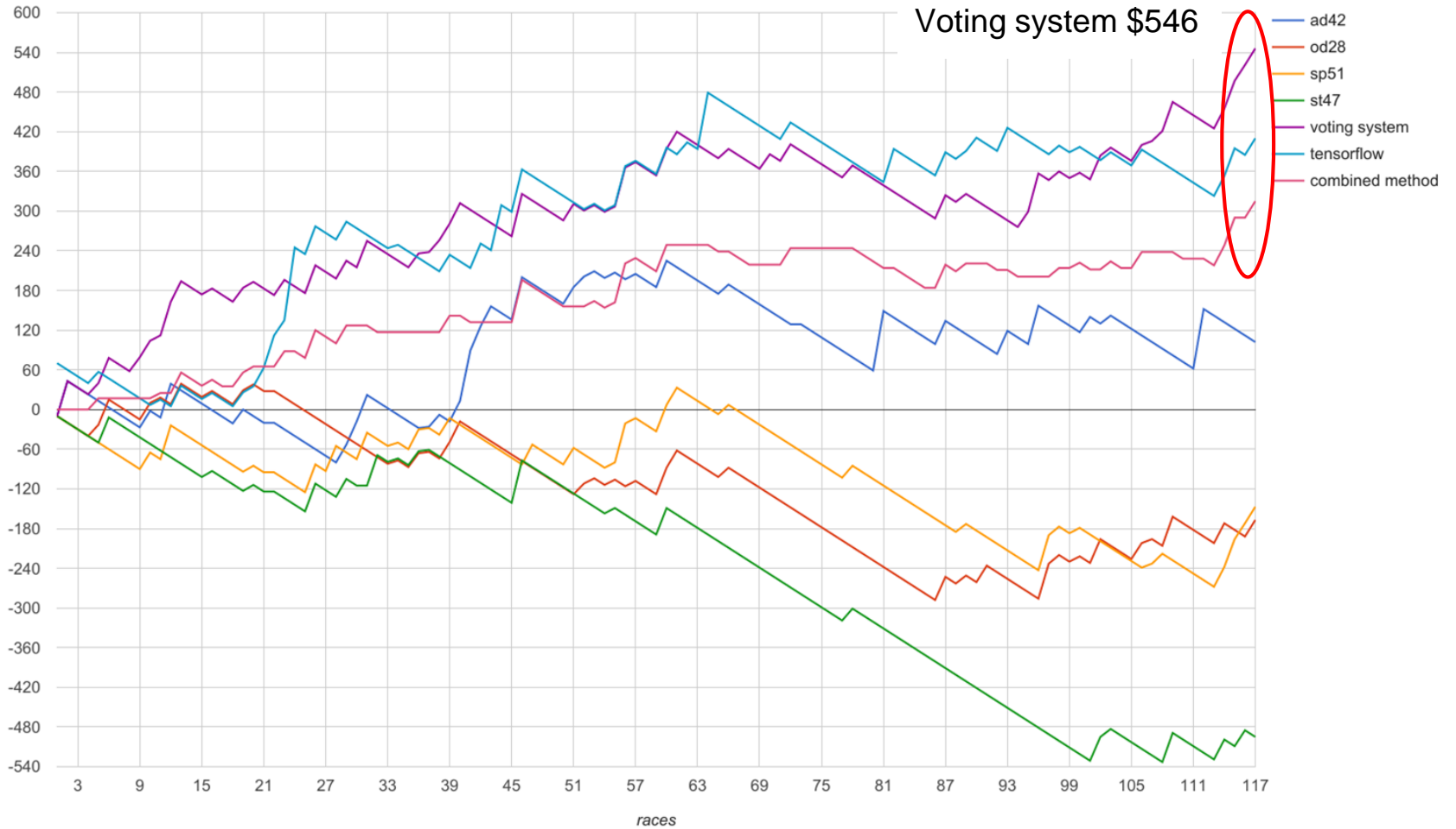
Choose one horse in each race, and buy 'win' for it

\$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

Cumulative net gain and loss analysis of '1 win'



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

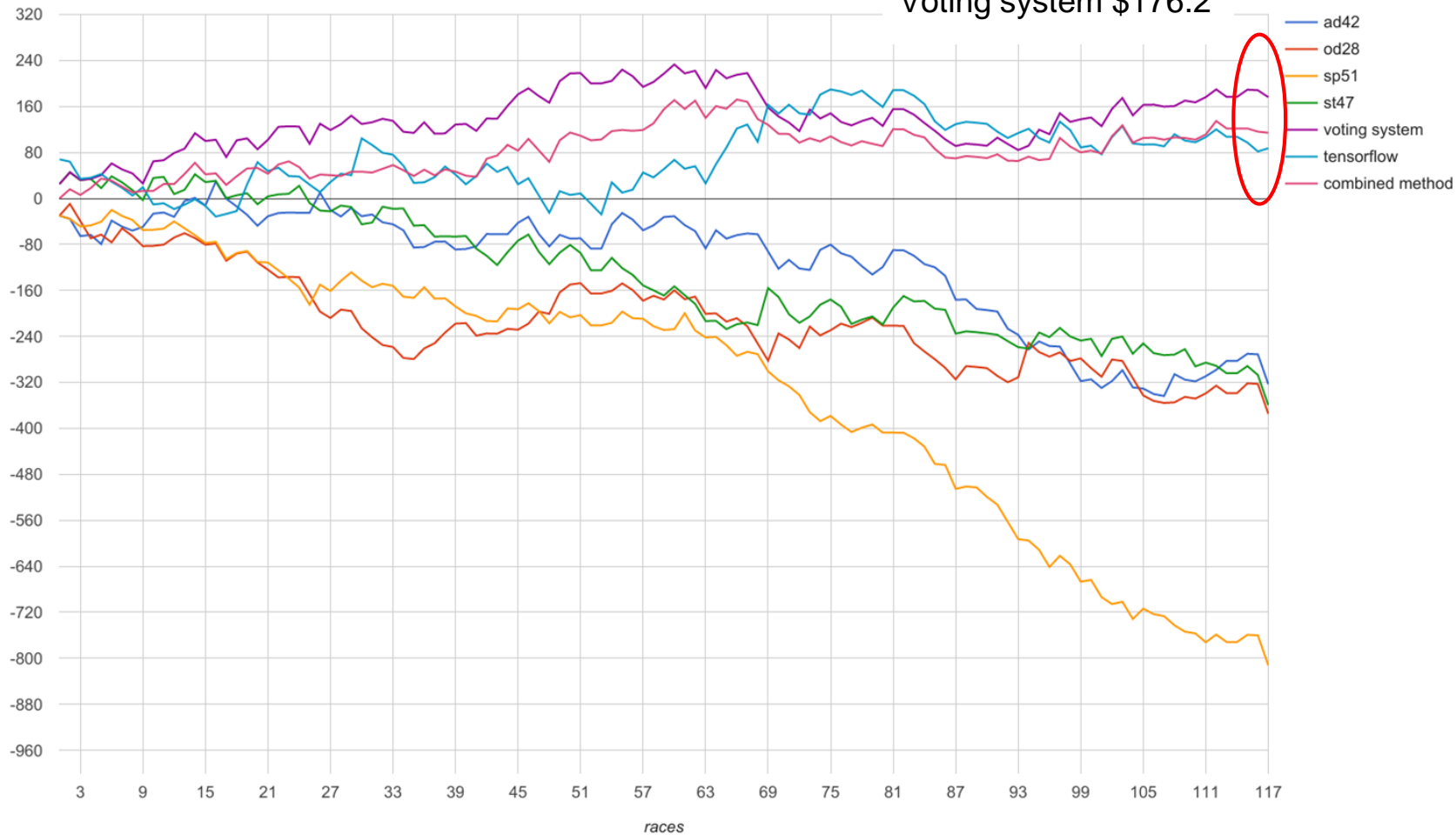
Total cost is \$3510 (Except Combined method)

Combined method:

only bet on 235 horses in 114 races over 117 races, total cost is \$2350

Cumulative net gain and loss analysis of '3 place'

Voting system \$176.2



Gain and loss analysis - 1 win without TensorFlow

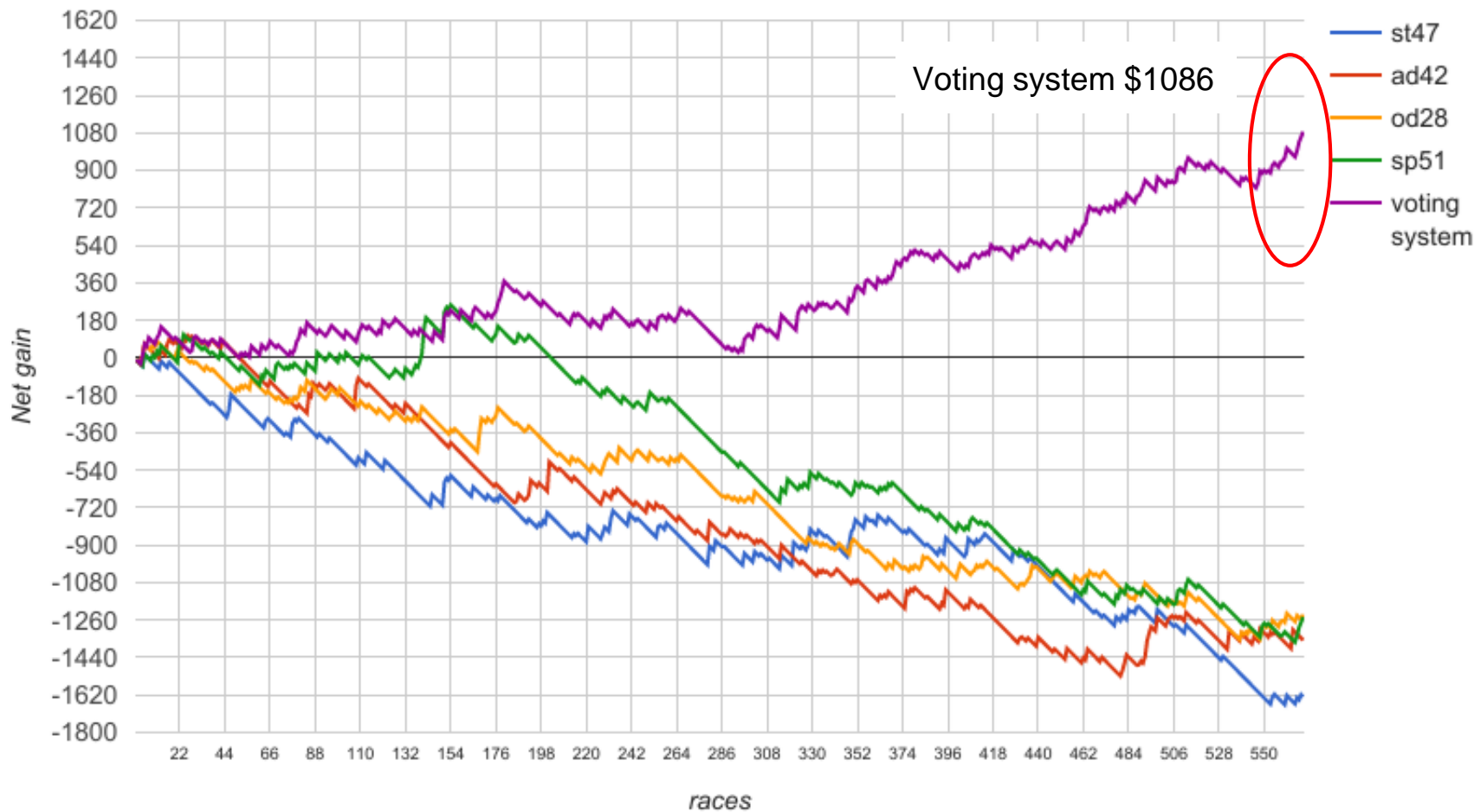
Choose one horse in each race, and buy 'win' for it

Total 569 races

\$10 for each horse

Total cost is \$5690

Cumulative net gain and loss analysis of '1 win' over 569 races



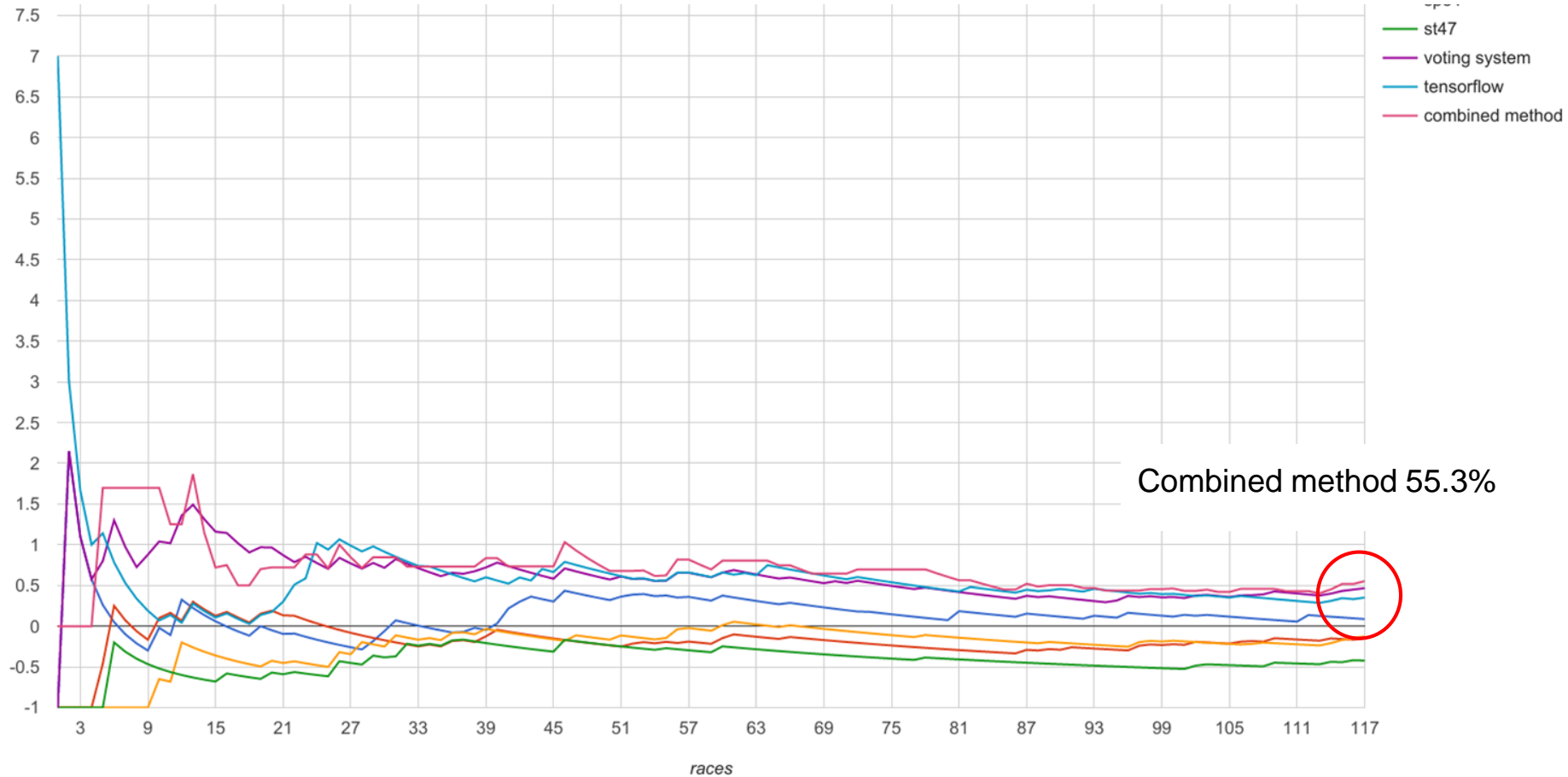
Profit and loss percentage

Combined method can't get highest net gain

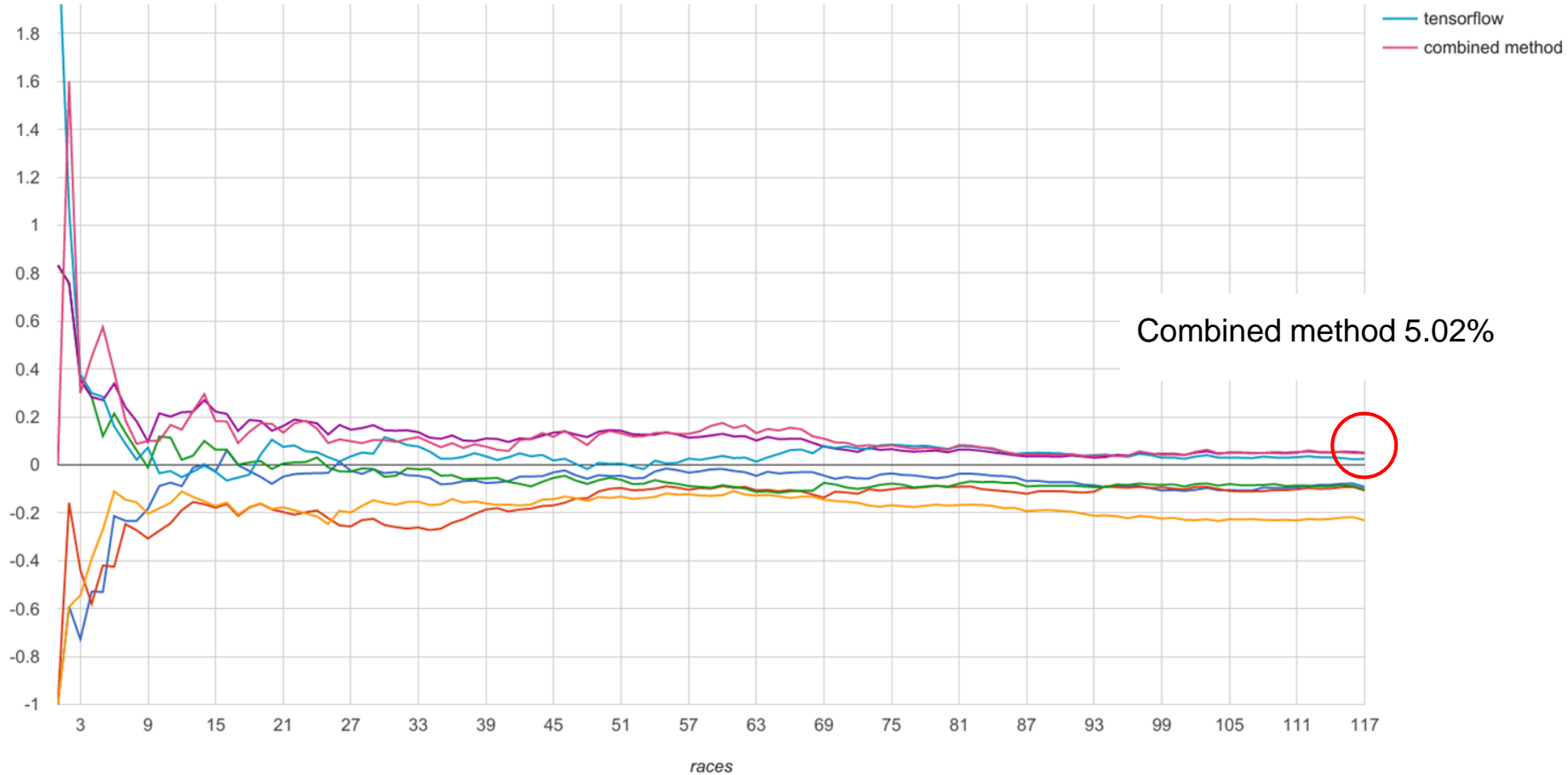
But it doesn't mean that combined method is not good

Profit and loss percentage = net gain over money spent

Profit and loss percentage - 1 win



Profit and loss percentage - 3 place



Maximization of returns

Combined method (highest profit percentage)

1 win

Further improvement:

Only bet on horses with win odds lower than 10

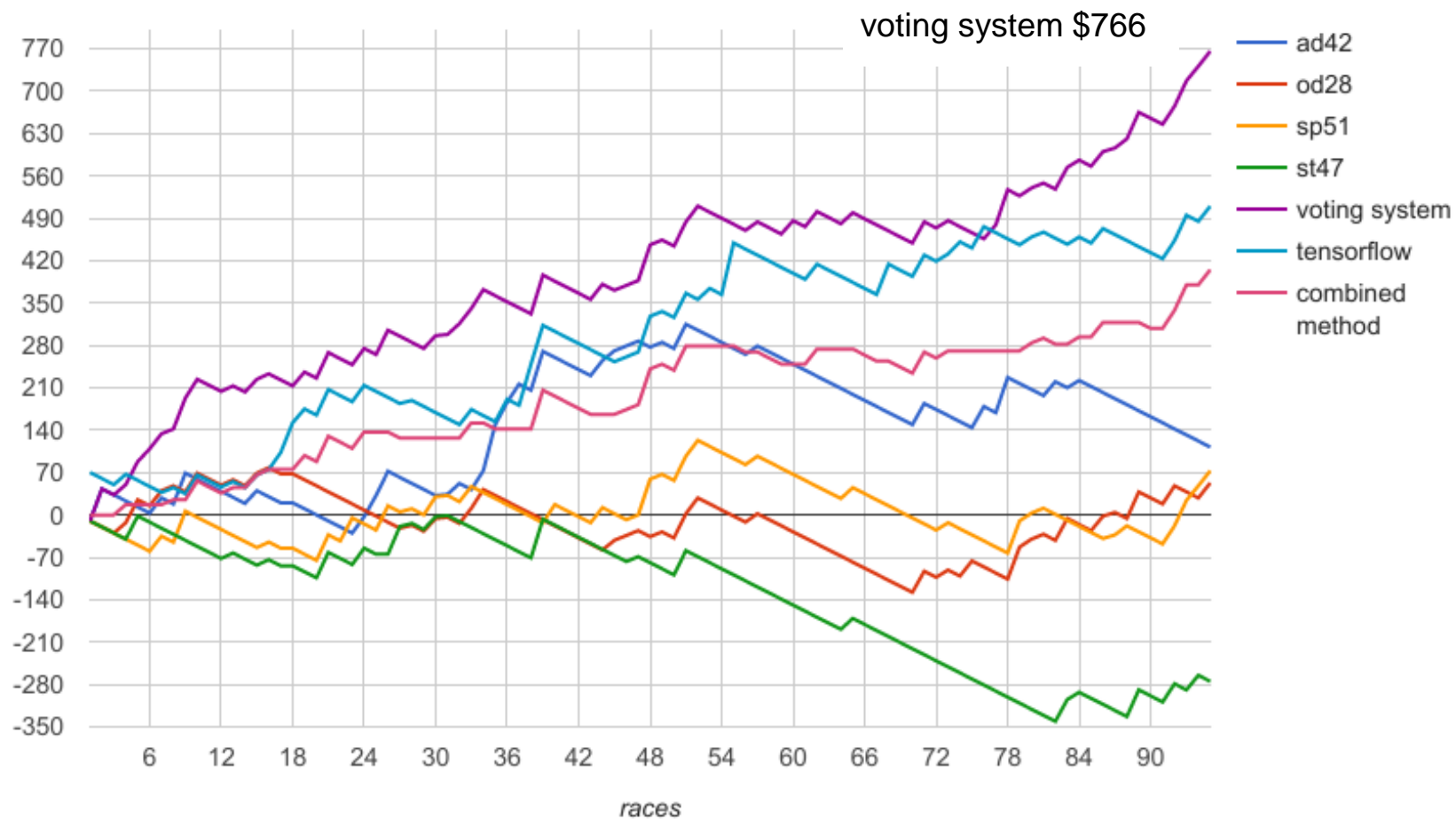
Maximization of returns - Result

1 win

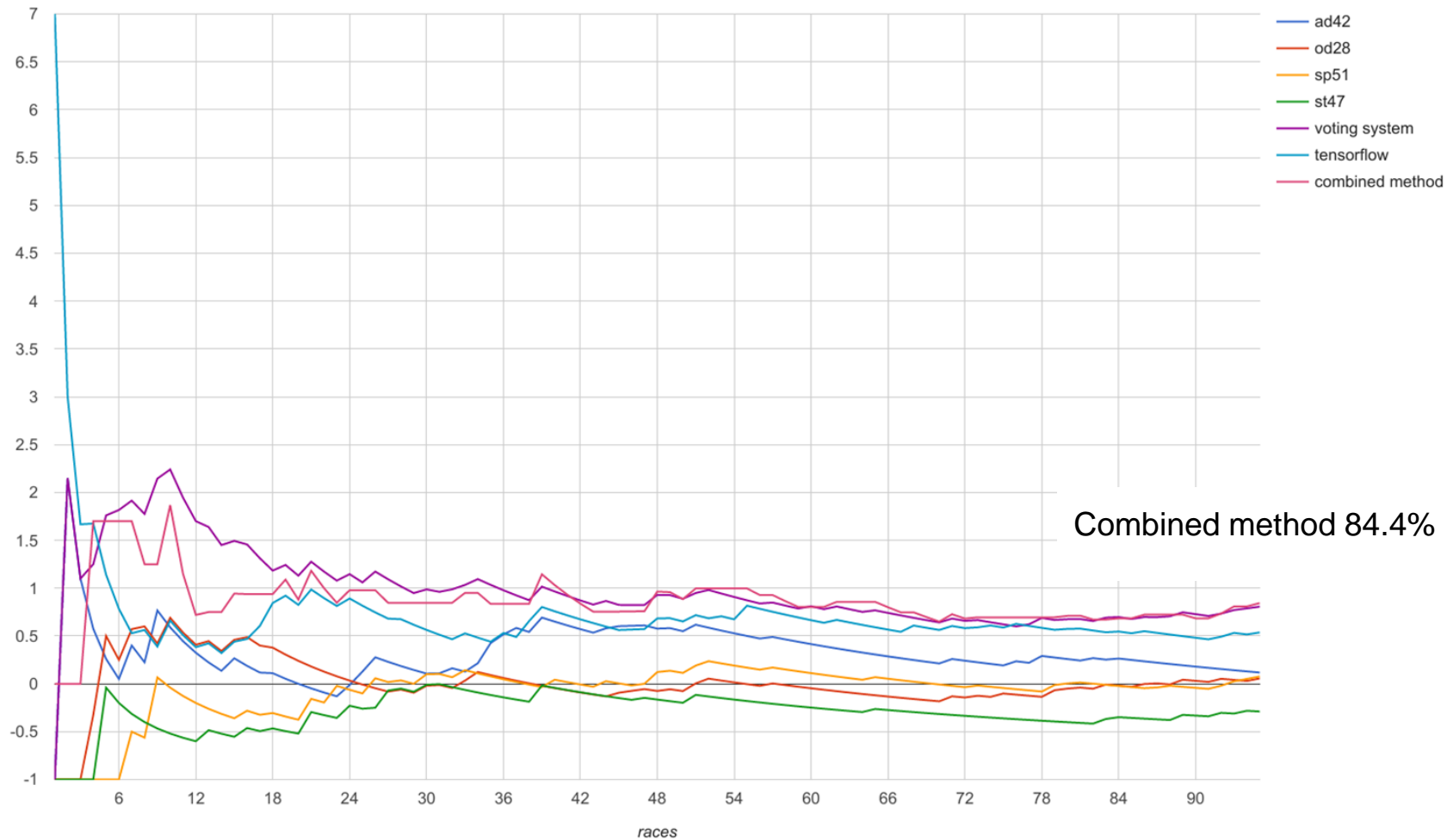
bet on 95 races over 117 races

For combined method: only bet on 48 races

Cumulative net gain of '1 win'(win odds<10)



Profit and loss percentage of '1 win'(win odds<10)



Combined method 84.4%

Maximization of returns - Result

Tips provider	Net gain		profit and loss percentage	
	Before	After	Before	After
ad42	102	112	8.72%	11.8%
od28	-167	53	-14.3%	5.58%
sp51	-147	73	-12.6%	7.68%
st47	-495	-275	-42.3%	-28.9%
voting system	546	766	46.7%	80.6%
TensorFlow	410	510	35.0%	53.7%
Combined method	315	405 (in 48 races)	55.3%	84.4% (in 48 races)

Let's see our prediction on 20-4-2017

race no	1st	2nd	3rd	predicted winner from voting system	Win Odds
1	5	2	10	5	2.5
2	4	2	8	3	5.7
3	7	9	4	9	4.7
4	1	7	12	1	4.6
5	2	10	5	2	4.6
6	1	8	2	1	4.7
7	8	4	2	8	5.6
8	3	12	9	7	16

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