



Base Strategy Boosting with ML

2019147006 박호찬

Contents

Topic

1. Topic

Base
Strategy

2. Base Strategy

Basic
Concept

3. Basic Concept

Methods

4. Methods

Result

5. Result

Topic

Topic

Base
Strategy

Basic
Concept

Methods

Result

Motivation

기본 트레이딩 전략의 성능을 향상시킬 수 있는 다양한 방안을 탐색하겠다.

Base Strategy

**Data
Pre-processing**

**Machine
Learning**

Base Strategy

Data

Kosdaq 150 Time bars

	A	B	C	D	
1	date_time	tick_num	open	high	
2	2021-01-04 10:00:01.899669	81	389.15	389.15	value
3	2021-01-04 10:00:03.907859	217	389.05	389.05	7.35
4	2021-01-04 10:00:08.327160	357	389.55	389.55	37.4
5	2021-01-04 10:00:12.859526	486	389.5	389.5	7.35
6	2021-01-04 10:00:21.659149	624	389.55	389.55	3.35
7	2021-01-04 10:00:24.959423	723	389.85	390.15	06.6
8	2021-01-04 10:00:28.823400	864	390.15	390.15	41.8
9	2021-01-04 10:00:34.588086	1015	390.25	390.25	1.05
10	2021-01-04 10:00:38.339601	1149	389.75	389.75	5.15
11	2021-01-04 10:00:44.131202	1288	389.1	389.1	40.3
12	2021-01-04 10:00:51.523688	1459	389.25	389.25	12.6
13	2021-01-04 10:01:00.187167	1613	389.15	389.15	9.75
14	2021-01-04 10:01:03.023159	1755	389.4	389.4	01.6
15	2021-01-04 10:01:07.555161	1884	389.75	389.75	845
16	2021-01-04 10:01:12.863176	2027	389.45	389.45	4.75
17	2021-01-04 10:01:18.371779	2119	389.65	389.65	35.7
18	2021-01-04 10:01:26.335787	2261	389.8	390.15	35.7
19	2021-01-04 10:01:35.319351	2412	389.75	389.75	35.7
20	2021-01-04 10:01:42.543167	2539	389.8	389.8	19.2
21	2021-01-04 10:01:52.279149	2673	389.55	389.55	42.8
22	2021-01-04 10:02:01.531209	2788	389.8	389.8	66.4
					6.25

KQ150 (1)

+

Kp200F_volume_bars_500

+

Base Strategy

Motivation

매도세가 약해지고 매수세가 강하게 이어졌을 때를 포착하여 매수를 하겠다.

$$f_T = \frac{C_T - O_t}{\max_{s \in [t, T]} H_s - \min_{s \in [t, T]} L_s}$$

C_T : Close Price at the end of window (종가)

O_t : Open Price at the start of window (시가)

$\max_{s \in [t, T]} H_s$: Highest Price in the window scope (고가)

$\min_{s \in [t, T]} L_s$: Lowest Price in the window scope (저가)

signal

$$= \min_{b \in \{s, m\}} \{ b: f_t \leq \text{thres and } f_{t+b} \geq 0 \}$$

s : scope length

m : max length (daily length)

Topic

Base
Strategy

Basic
Concept

Methods

Result

Base Strategy

Motivation

Topic

Base
Strategy

Basic
Concept

Methods

Result



Base Strategy

이상 시그널 발생 case1



-> parkinson volatility를 이용하여 변동성이 낮은 구간에서의 진입 시그널 제거

case2



-> 기준치 아래로 local minimum을 2번 찍은 후 진입

Topic

Base
Strategy

Basic
Concept

Methods

Result

Topic

Base
Strategy

Basic
Concept

Methods

Result

Parkinson Volatility:

$$\sigma_{parkinson} = \sqrt{\frac{1}{4 \ln(2) * n} \sum_{i=1}^n \ln \left(\frac{H_i}{L_i} \right)^2}$$

→ 변동성이 낮은 구간에서 진입 시그널을 제거

Base Strategy

이상 시그널 발생 case1



-> parkinson volatility를 이용하여 변동성이 낮은 구간에서의 진입 시그널 제거

case2



-> 기준치 아래로 local minimum을 2번 찍은 후 진입

Topic

Base
Strategy

Basic
Concept

Methods

Result

Base Strategy

Topic

Base
Strategy

Basic
Concept

Methods

Result

전략 평가 방법: $pnl = (\text{청산시점} - \text{진입시점} - \text{slippage}) / \text{진입시점}$

```
2021-11-11 10:21:00 2021-11-11 10:23:00 2021-11-11 10:26:00  
0.9993066561846508  
Cumulative PNL: 1.4965101292438205  
Process 1440 Done
```

```
Sharpe: 7.008737819704887  
Total Entrance: 950  
Win Ratio: 0.5889387144992526
```

기본 전략 - 샤프비율: 7.01, 승률: 59%

Base Strategy

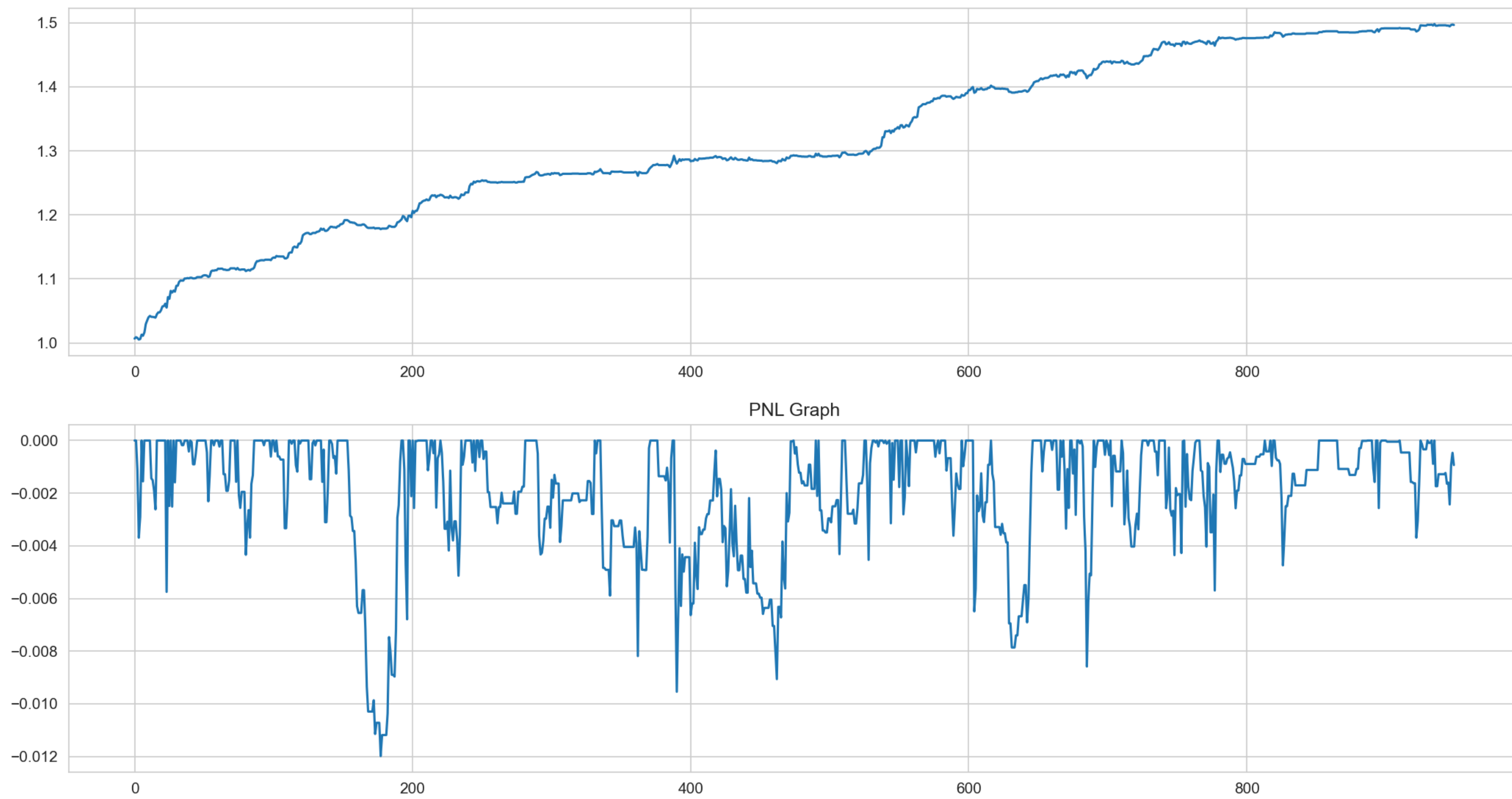
Topic

Base
Strategy

Basic
Concept

Methods

Result



Basic Concept

Triple Barrier Method



Topic

Base
Strategy

Basic
Concept

Methods

Result

Basic Concept

Topic

Meta-Labeling Method

Base
Strategy

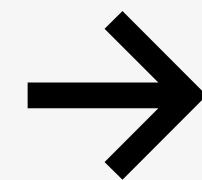
: 생성한 시그널이 진짜 수익률을 가져다 줄 수 있는지 이차검증

Basic
Concept

<실제 수익률>

음수(-)

양수(+)



< meta label >

0 (가짜 시그널)

1 (진짜 시그널)

Methods

Result

Methods

기본 전략을 통해 얻은 매수/매도 시그널을 이용하여 진입(1), 손절(-1) 값 삽입

	A	B	C	D	E	F
1	Date-Time	Open	High	Low	Close	side
137960	2018-01-03 14:07	1447.9	1450.4	1447.9	1450.4	1
137963	2018-01-03 14:10	1449.9	1454.4	1449.9	1453.4	-1
138006	2018-01-03 14:53	1455.2	1457.6	1455.2	1457.6	1
138009	2018-01-03 14:56	1459.2	1460.6	1458.8	1460.6	-1
138069	2018-01-04 9:20	1459.6	1459.9	1456.7	1459.9	1
138072	2018-01-04 9:23	1463.1	1463.4	1461.8	1463.4	-1
138181	2018-01-04 11:12	1456.9	1456.9	1456.2	1456.9	1
138184	2018-01-04 11:15	1458.2	1458.5	1458	1458.5	-1
138428	2018-01-04 15:19	1422.6	1423.5	1422.2	1423.5	1
138431	2018-01-04 15:22	1421.1	1421.8	1421.1	1421.8	-1
138519	2018-01-05 10:14	1446.3	1446.5	1444.2	1446.5	1
138522	2018-01-05 10:17	1444.4	1444.5	1443.3	1444.5	-1

Methods

Topic

Triple barrier labeling

Base
Strategy

수익률을 내는 구간 샘플링 (t1: vertical barrier)
& 예상 수익률 (trgt)

Basic
Concept

Methods

Result

	A	B	C
1	t1	trgt	side
14355	2018-01-03 14:10	0.001389	1
14358	2018-01-03 14:11	0.001356	-1
14401	2018-01-03 14:55	0.000768	1
14404	2018-01-03 14:58	0.000886	-1
14456	2018-01-04 9:21	0.001251	1
14459	2018-01-04 9:25	0.001477	-1
14540	2018-01-04 11:15	0.000552	1
14543	2018-01-04 11:16	0.000579	-1
14653	2018-01-04 15:23	0.001595	1
14656	2018-01-04 15:25	0.00148	-1
14742	2018-01-05 10:15	0.0011	1
14745	2018-01-05 10:20	0.001092	-1
14809	2018-01-05 13:35	0.000881	1
14812	2018-01-05 13:41	0.001022	-1
15060	2018-01-08 13:41	0.001894	1
15062	2018-01-08 13:54	0.001924	-1
15229	2018-01-09 10:03	0.000642	1

Methods

Meta labeling

시그널에 따라 실제 수익률 계산했을 때,
수익률이 음수면 가짜 시그널(0),
양수면 진짜 시그널(1)로 표시.

	A	B	C
1		ret	bin
14355	2018-01-03 14:07	0.002068	1
14358	2018-01-03 14:10	-0.00296	0
14401	2018-01-03 14:15	0.001917	1
14404	2018-01-03 14:16	0.001917	1
14456	2018-01-04 9:51	-0.00157	0
14459	2018-01-04 9:23	-0.00157	0
14540	2018-01-04 11:12	0.001098	1
14543	2018-01-04 11:13	-0.00225	0
14653	2018-01-04 15:09	-0.00225	0
14656	2018-01-04 15:10	-0.00225	0
14742	2018-01-05 10:14	0.001177	1
14745	2018-01-05 10:17	0.001246	1
14809	2018-01-05 13:31	0.000893	1
14812	2018-01-05 13:34	0.001099	1
15060	2018-01-08 13:39	-0.00193	0
15062	2018-01-08 13:41	-0.00227	0
15229	2018-01-09 9:59	-0.00066	0

Result

Topic

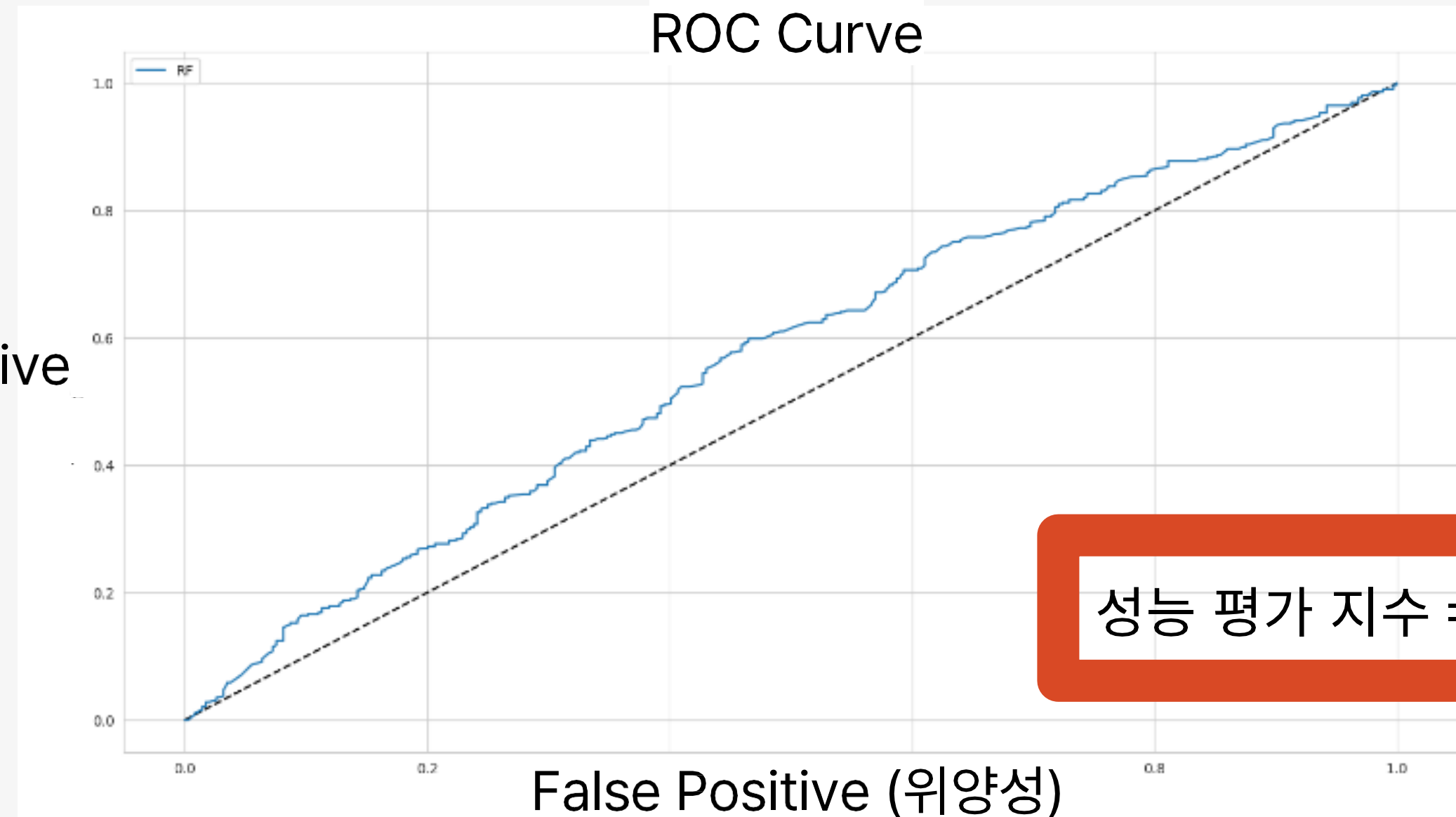
Base
Strategy

Basic
Concept

Methods

Result

True Positive
(진양성)



성능 평가 지수 = $TP / (TP+FP)$

TP = True Positive = 수익률이 오를 것이라 예상해서 진입했을 때 실제로 오를 확률
FP = False Positive = 수익률이 오를 것이라 예상해서 진입했지만 실제로 내릴 확률

Result

Topic

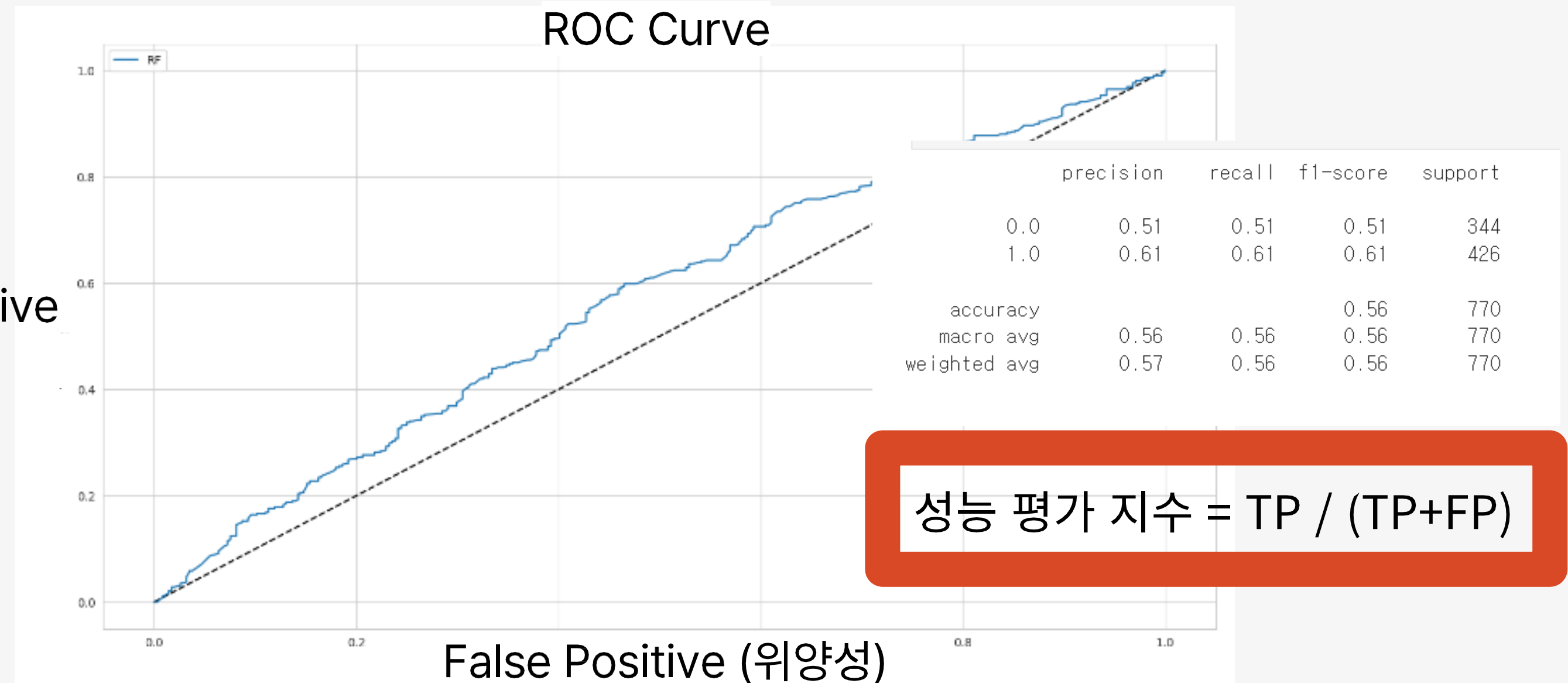
Base
Strategy

Basic
Concept

Methods

Result

True Positive
(진양성)



TP = True Positive = 수익률이 오를 것이라 예상해서 진입했을 때 실제로 오를 확률
FP = False Positive = 수익률이 오를 것이라 예상해서 진입했지만 실제로 내릴 확률

Thank You

CONTACT

john1408@yonsei.ac.kr