R.Dash 논문리뷰

19.02.27

김도완

Ke Ai
ADVANCING RESEARCH
EVOLVING SCIENCE

CrossMark

Available online at www.sciencedirect.com

ScienceDirect

The Journal of Finance and Data Science 2 (2016) 42–57

http://www.keaipublishing.com/en/journals/

A hybrid stock trading framework integrating technical analysis with machine learning techniques

Rajashree Dash [a,*,1], Pradipta Kishore Dash [b,2]

[a] *Computer Science & Engineering Department, ITER, Siksha 'O' Anusandhan University, Bhubaneswar, Odisha, 751030, India*
[b] *Siksha 'O' Anusandhan University, Bhubaneswar, Odisha, 751030, India*

Received 30 December 2015; revised 3 March 2016; accepted 8 March 2016
Available online 22 March 2016

머신러닝을 이용, 기술적지표 분석을 통한 주가트레이딩 프레임워크
R. Dash, 2016

1. 기술적 지표 추가

날짜	시가	고가	저가	종가	거래량	MA15	SlowK	SlowD	MACD	MACDsign	MACDhist	RSI14	WILR14
20120220	226500	228000	223500	224000	560653.0	219300.000000	71.296182	64.246716	-906.647208	-1509.717851	603.070644	57.162524	-22.857143
20120221	225000	227000	223500	226500	431118.0	219666.666667	82.232143	72.500342	-404.897906	-1288.753862	883.855956	59.876635	-8.571429
20120222	227500	228500	221500	224000	449186.0	220033.333333	80.432900	77.987075	-206.605457	-1072.324181	865.718724	56.052086	-25.000000
20120223	224500	224500	217500	219000	553306.0	220100.000000	64.595158	75.753400	-447.754047	-947.410154	499.656108	49.273337	-52.777778
20120224	216000	218000	212500	216500	861826.0	220100.000000	42.452301	62.493453	-831.015716	-924.131267	93.115551	46.260844	-72.727273
20120227	214500	215000	207000	210000	1023577.0	219933.333333	22.861039	43.302833	-1640.340206	-1067.373054	-572.967151	39.499093	-86.046512
20120228	207000	210000	205000	208500	818030.0	219233.333333	17.949035	27.754125	-2375.391129	-1328.976669	-1046.414460	38.114573	-85.106383
20120229	209000	216000	209000	216000	761238.0	218866.666667	25.218539	22.009538	-2325.925144	-1528.366364	-797.558780	47.940372	-53.191489
20120302	222000	223500	220000	220000	1030467.0	218866.666667	41.843972	28.337182	-1941.575155	-1611.008122	-330.567032	52.291104	-36.170213
20120305	220500	223000	218000	220000	478789.0	219033.333333	58.156028	41.739513	-1618.320033	-1612.470504	-5.849528	52.291104	-36.170213

기술적지표 추가 : TA-Lib 패키지 사용

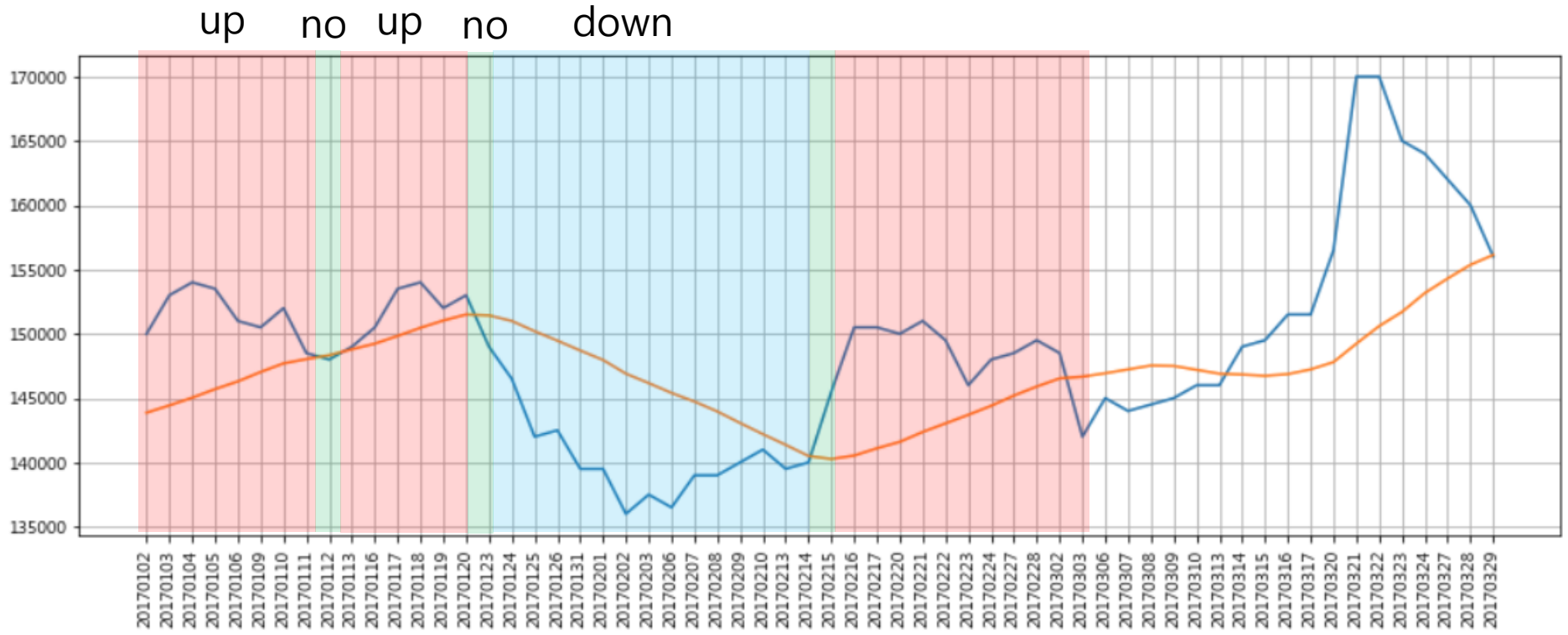
생성된 값 검증, HTS와 비교해봐야
더 다양한 지표를 만들어서 테스트해본다.
거래량과 관련된 지표, 외국인/기관 수급 등

2. Trend(추세)값 생성

날짜	종가	MA15	MA15_P1	MA15_P2	MA15_P3	MA15_P4	MA15_P5	Trend
20120227	210000	219933.333333	220100.000000	220100.000000	220033.333333	219666.666667	219300.000000	no
20120228	208500	219233.333333	219933.333333	220100.000000	220100.000000	220033.333333	219666.666667	down
20120229	216000	218866.666667	219233.333333	219933.333333	220100.000000	220100.000000	220033.333333	down
20120302	220000	218866.666667	218866.666667	219233.333333	219933.333333	220100.000000	220100.000000	no
20120305	220000	219033.333333	218866.666667	218866.666667	219233.333333	219933.333333	220100.000000	no
20120306	216500	218700.000000	219033.333333	218866.666667	218866.666667	219233.333333	219933.333333	down
20120307	213000	218300.000000	218700.000000	219033.333333	218866.666667	218866.666667	219233.333333	down
20120308	212000	217866.666667	218300.000000	218700.000000	219033.333333	218866.666667	218866.666667	down
20120309	212500	217433.333333	217866.666667	218300.000000	218700.000000	219033.333333	218866.666667	down
20120312	215500	216933.333333	217433.333333	217866.666667	218300.000000	218700.000000	219033.333333	down

MA15의 지난5일간을 확인하여 트렌드값을 만든다.

트렌드 up의 조건 : MA15가 지난5일간 최대일때 And 현종가 > MA15
트렌드 down의 조건 : MA15가 지난5일간 최소일때 And 현종가 < MA15
그 외 no


up
no
up
no
down
170000
165000
160000
155000
150000
145000
140000
135000
20170102
20170103
20170104
20170105
20170106
20170109
20170110
20170111
20170112
20170113
20170116
20170117
20170118
20170119
20170120
20170123
20170124
20170125
20170126
20170131
20170201
20170202
20170203
20170206
20170207
20170208
20170209
20170210
20170213
20170214
20170215
20170216
20170217
20170220
20170221
20170222
20170223
20170224
20170227
20170228
20170302
20170303
20170306
20170307
20170308
20170309
20170310
20170313
20170314
20170315
20170316
20170317
20170320
20170321
20170322
20170323
20170324
20170327
20170328
20170329

3. Trading signal 생성

날짜	종가	Trend	종가_F1	종가_F2	Trading signal
20170102	150000	up	153000.0	154000.0	1.000000
20170103	153000	up	154000.0	153500.0	1.000000
20170104	154000	up	153500.0	151000.0	0.500000
20170105	153500	up	151000.0	150500.0	0.500000
20170106	151000	up	150500.0	152000.0	0.833333
20170109	150500	up	152000.0	148500.0	0.714286
20170110	152000	up	148500.0	148000.0	0.500000
20170111	148500	up	148000.0	149000.0	0.750000
20170112	148000	no	149000.0	150500.0	0.500000
20170113	149000	up	150500.0	153500.0	1.000000
20170116	150500	up	153500.0	154000.0	1.000000
20170117	153500	up	154000.0	152000.0	0.625000
20170118	154000	up	152000.0	153000.0	0.500000
20170119	152000	up	153000.0	149000.0	0.625000
20170120	153000	up	149000.0	146500.0	0.500000
20170123	149000	no	146500.0	142000.0	0.000000
20170124	146500	no	142000.0	142500.0	0.000000
20170125	142000	down	142500.0	139500.0	0.083333
20170126	142500	down	139500.0	139500.0	0.000000
20170131	139500	down	139500.0	136000.0	0.000000

Trend값과 미래종가를 이용해
Trading signal을 만든다.

(현종가, +1종가, +2종가)
max_cp최대/ min_cp최소치 확인

Trend가 up일때
(max_cp – 현종가) / (max_cp – min_cp)*0.5+0.5

Trend가 down 또는 no일때
(max_cp – 현종가) / (max_cp – min_cp) *0.5

현종가가 미래종가보다 작을때는 1~0.5
현종가가 미래종가보다 클때는 0.5~0

For up trend:

$$Tr_i = \frac{[cp_i - \min cp]}{[\max cp - \min cp]} \times 0.5 + 0.5$$
$$\min cp = \min(cp_i,\ cp_{i+1},\ cp_{i+2}) \tag{11}$$
$$\max cp = \max(cp_i,\ cp_{i+1},\ cp_{i+2})$$

For down trend:

$$Tr_i = \frac{[cp_i - \min cp]}{[\max cp - \min cp]} \times 0.5 \tag{12}$$

Where cpi, cpi+1, cpi+2, are the closing price of the ith, (i+1)th, (i+2)th trading days respectively.

Table 2 represents the trading signal generated from the trend analysis for the sample data set using the equations (11) and (12).

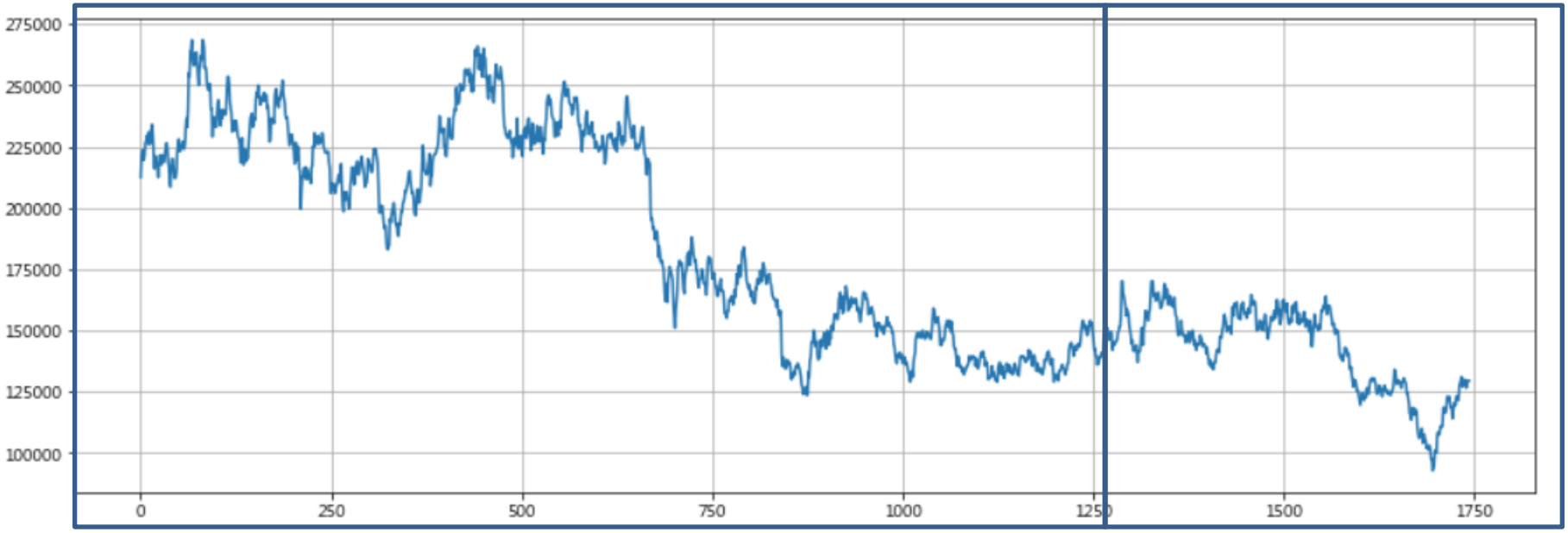
4. Feature(지표) 선택 및 스케일링

feature 10개 → 목표값 target

날짜	종가	거래량	MA15	SlowK	SlowD	MACD	MACDsign	MACDhist	RSI14	WILR14	Trading signal
20120227	0.667046	0.205164	0.743971	0.224361	0.435286	0.442448	0.487085	0.394478	0.354799	0.139535	0.4
20120228	0.658509	0.156096	0.739635	0.174119	0.270449	0.410613	0.474609	0.334512	0.333074	0.148936	0.5
20120229	0.701195	0.142539	0.737364	0.248475	0.209548	0.412755	0.465100	0.366031	0.487252	0.468085	0.5
20120302	0.723961	0.206809	0.737364	0.418527	0.276630	0.429401	0.461159	0.425180	0.555520	0.638298	0.0
20120305	0.723961	0.075114	0.738396	0.585373	0.418713	0.443401	0.461089	0.466308	0.555520	0.638298	0.0

MinMax Scale

5. 학습용, 테스트용 데이터 분리



Train Set: 7

Test Set: 3

6. 모델 학습

```
from sklearn.ensemble import RandomForestRegressor

rf= RandomForestRegressor(n_estimators=100, random_state=0)
rf.fit(train_X, train_Y)
print('훈련 세트 정확도: {:.3f}'.format(rf.score(train_X, train_Y)))
print('테스트 세트 정확도: {:.3f}'.format(rf.score(test_X, test_Y)))
```

```
훈련 세트 정확도: 0.913
테스트 세트 정확도: 0.340
```

```
np.random.seed(5)

keras.backend.clear_session()
model = keras.models.Sequential()

model.add( layers.Dense(14, input_shape=(train_X.shape[1],), activation='relu' ))
model.add( layers.Dense(10, activation='relu' ))
model.add( layers.Dense(5, activation='relu' ))
model.add( layers.Dense(1) )

adam = optimizers.Adam(lr=0.01, decay=0.0001)
model.compile(loss = 'mse', optimizer = 'adam', metrics =['accuracy'])
model.summary()
```

훈련세트, 테스트세트 정확도 : 0.36~0.40

6. 모델 학습

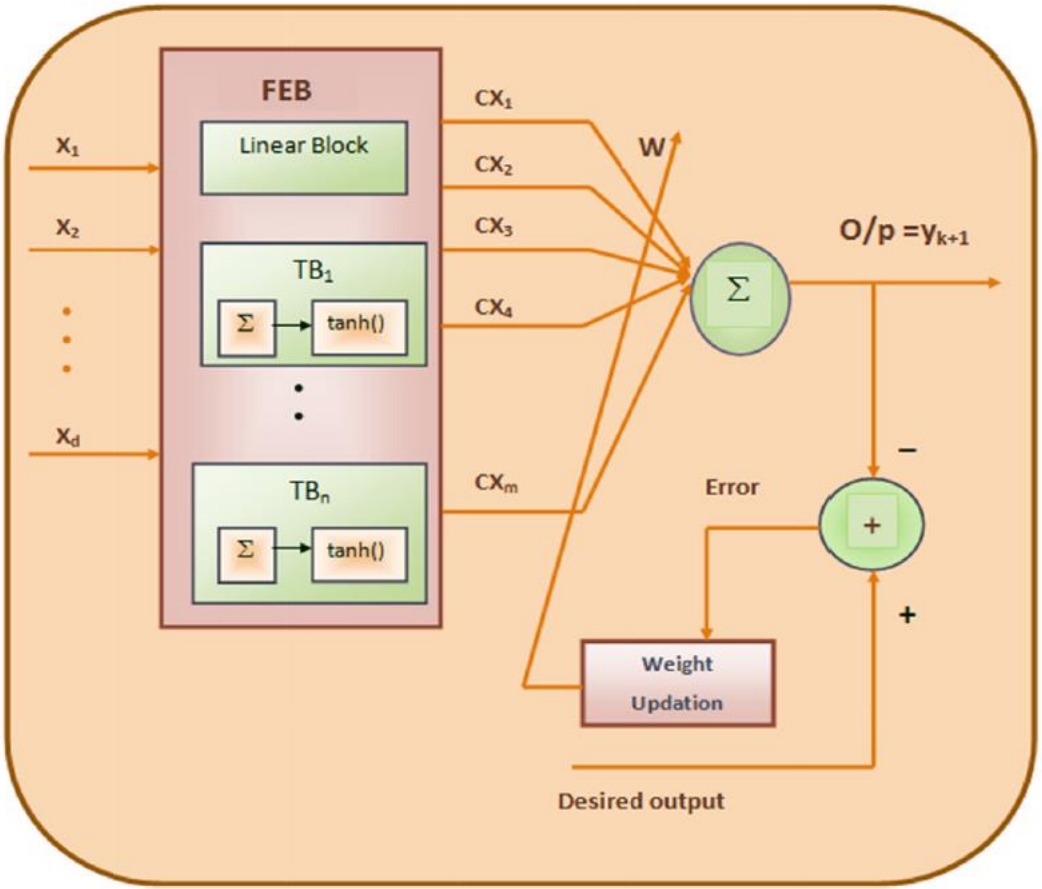


Fig. 1. Architecture of computational efficient FLANN (CEFLANN).

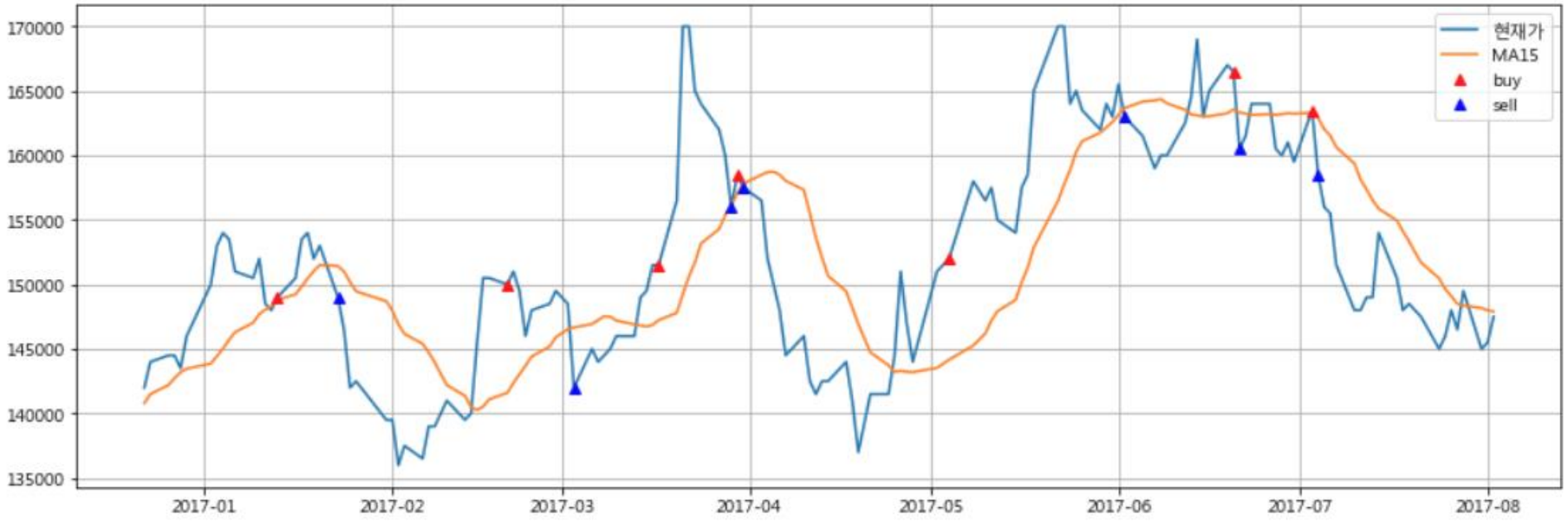
6. 예측값 -> Buy and Sell 신호 생성

날짜	pred	Trend2	TS
20170106	0.690380	1	0
20170109	0.709813	1	0
20170110	0.723189	1	0
20170111	0.692525	1	0
20170112	0.179729	-1	0
20170113	0.699338	1	1
20170116	0.743258	1	0
20170117	0.722721	1	0
20170118	0.687939	1	0
20170119	0.693636	1	0
20170120	0.716894	1	0
20170123	0.157038	-1	-1

(1) 예측값(pred) -> Trend2
0.5보다 크면 1 작으면 -1

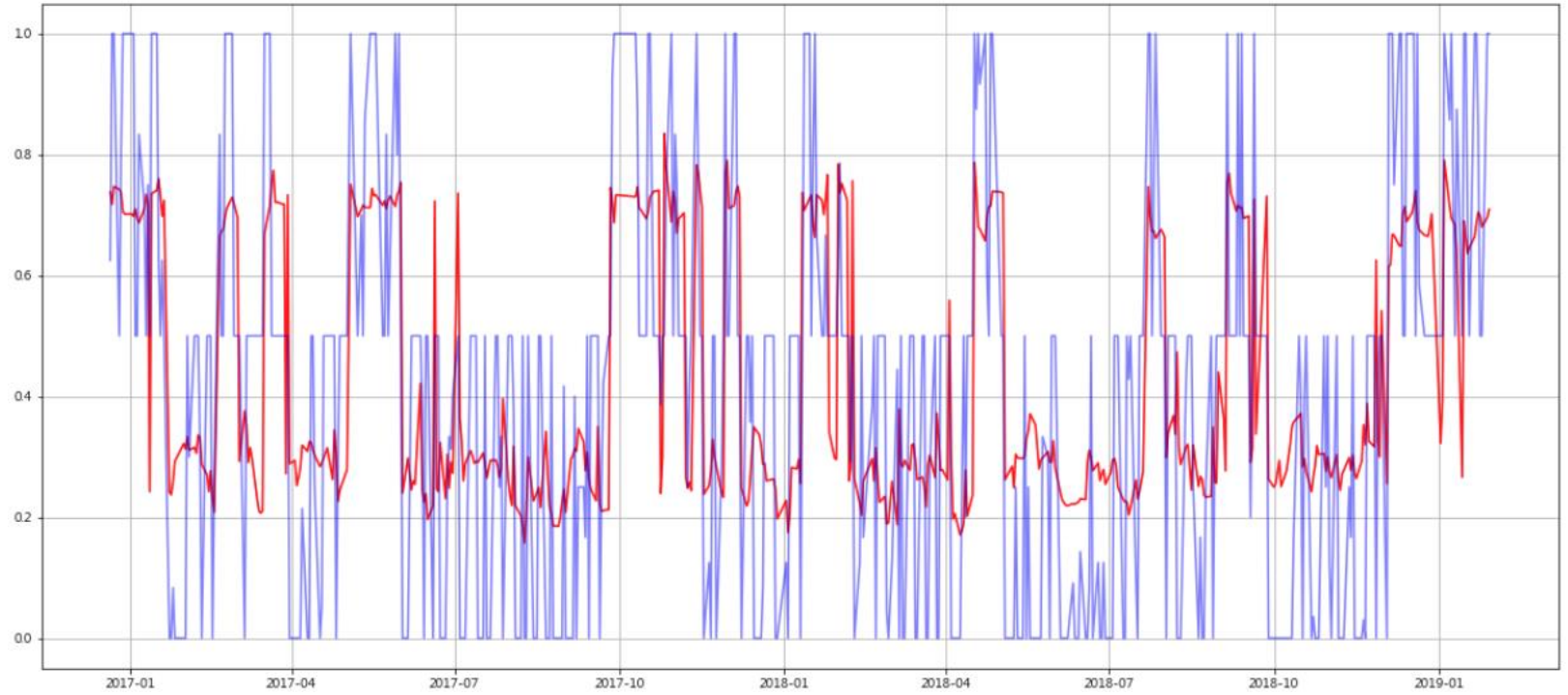
(2) Trend2가 변화할때 TS값
-1->1 buy 신호
1->-1 sell 신호

7. 백테스팅


현재가
MA15
buy
sell
170000
165000
160000
155000
150000
145000
140000
135000
2017-01
2017-02
2017-03
2017-04
2017-05
2017-06
2017-07
2017-08

Trading signal 값 비교
실제값 vs 예측값

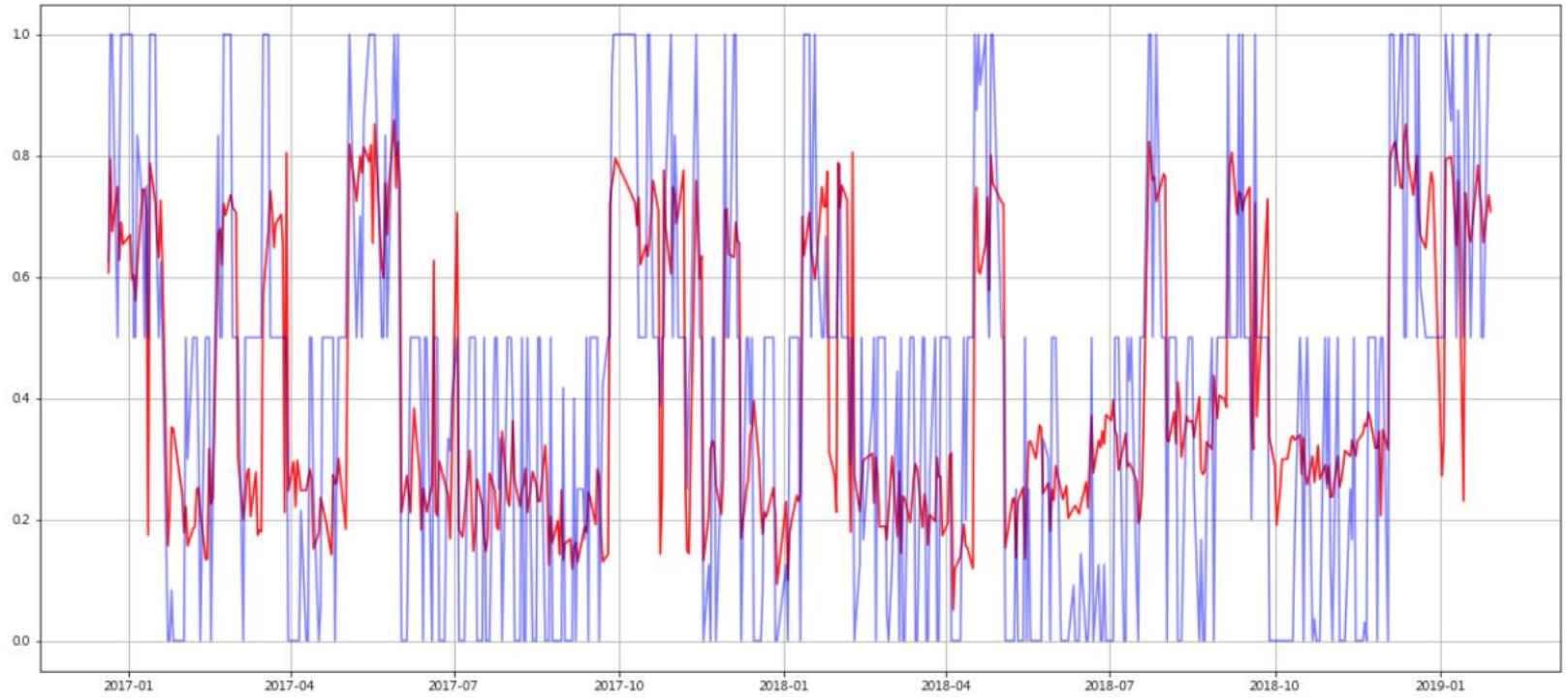
DNN1


1.0
0.8
0.6
0.4
0.2
0.0
2017-01
2017-04
2017-07
2017-10
2018-01
2018-04
2018-07
2018-10
2019-01

DNN2


1.0
0.8
0.6
0.4
0.2
0.0
2017-01
2017-04
2017-07
2017-10
2018-01
2018-04
2018-07
2018-10
2019-01

RF

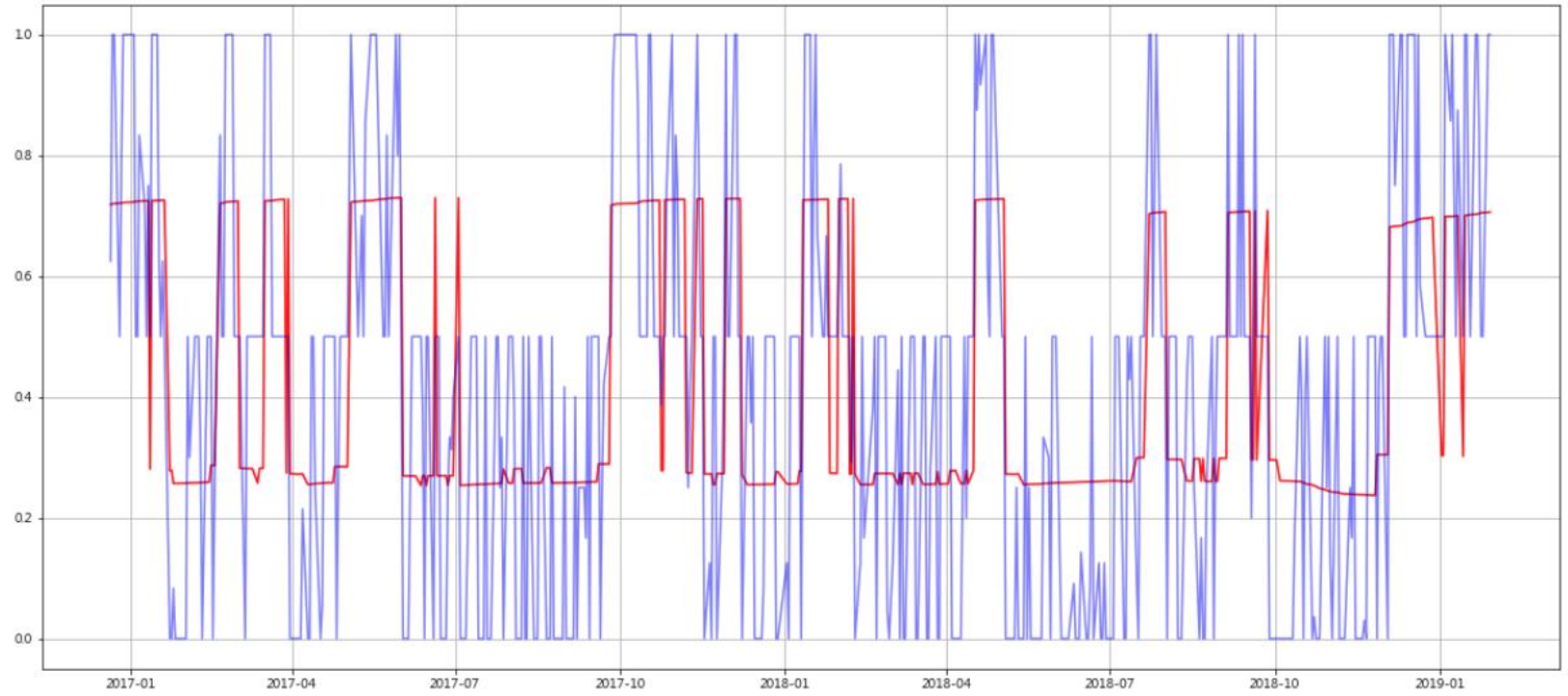

1.0
0.8
0.6
0.4
0.2
0.0
2017-01
2017-04
2017-07
2017-10
2018-01
2018-04
2018-07
2018-10
2019-01

DNN3 (MA5, MADiff 5-15)

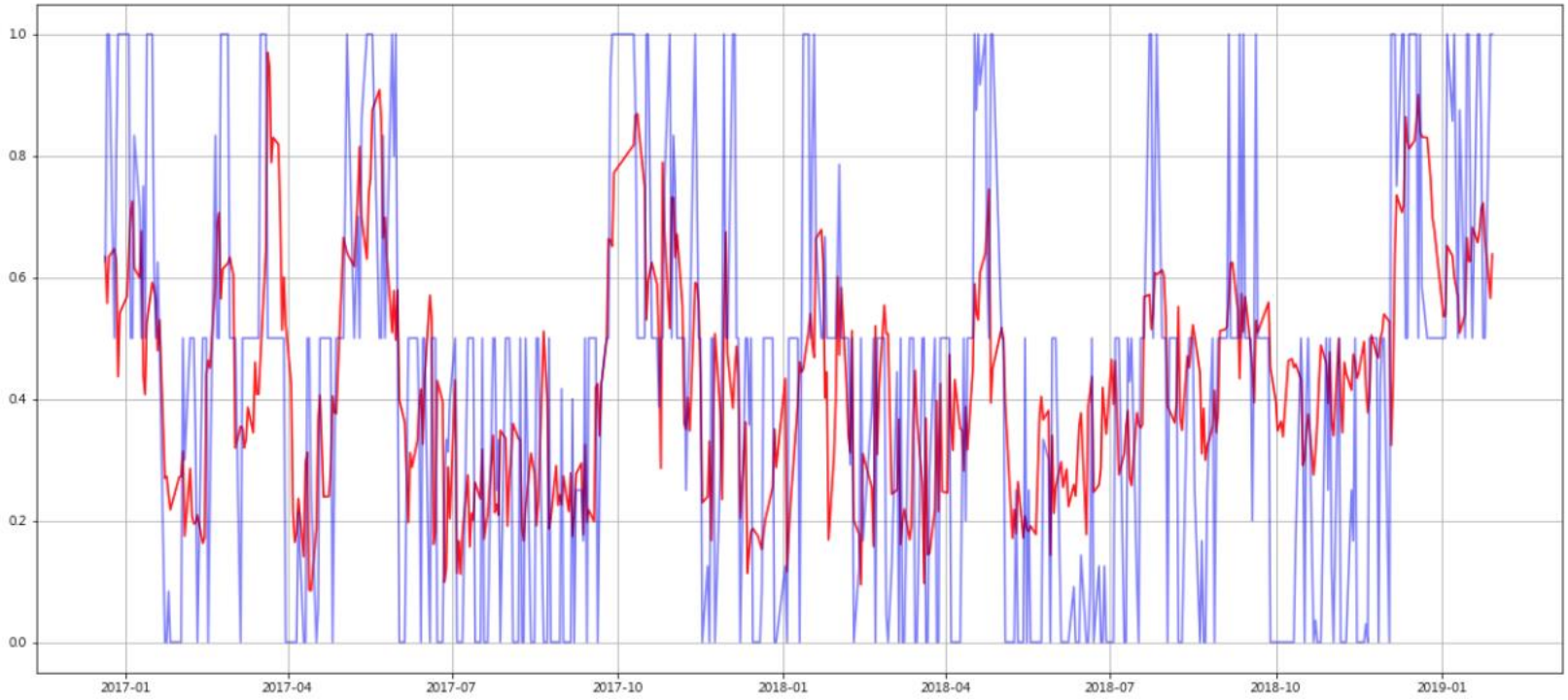
DNN3 (MADiff 5-20)

20170112, 0620, 0703, 20180126/29/30

DNN3 (트렌드만)


1.0
0.8
0.6
0.4
0.2
0.0
2017-01
2017-04
2017-07
2017-10
2018-01
2018-04
2018-07
2018-10
2019-01

DNN3 (트렌드제외)


1.0
0.8
0.6
0.4
0.2
0.0
2017-01
2017-04
2017-07
2017-10
2018-01
2018-04
2018-07
2018-10
2019-01