1. **Feature Variables**

삼성전자의 일별 주가 데이터(open, high, low, close, volume, adjusted, 2007년 1월 2일 ~ 2020년 10월 16일)에 기반하여 생성된 여러 가지 기술적 지표 변수들.

**ROC** : Rate of Change (the percentage difference of a series over two observations) over 1 period

**ROC5** : Rate of Change over 5 period

**Momemtum2** : 2-day momentum over 2 periods.

**BBmavg** : Variable related to Bollinger Bands. The middle Moving Average (see notes).

**BBup** : Variable related to Bollinger Bands. The upper Bollinger Band.

**BBpctB** : Variable related to Bollinger Bands. The %B calculation.

**DIp** : Variable related to Directional Movement Index. The positive Direction Index.

**DIn** : Variable related to Directional Movement Index. The negative Direction Index.

**DI** : Variable related to Directional Movement Index. The Direction Index.

**ADX** : Variable related to Directional Movement Index. The Average Direction Index (trend strength).

**ema20** : Exponential moving average over the past n(=20) observations.

**sma20** : Simple moving average over the past n(=20) observations.

**ema60** : Exponential moving average over the past n(=60) observations.

**sma60** : Simple moving average over the past n(=60) observations.

**MACDoscil** : Variable related to MACD. The price oscillator.

**MACDsignal** : Variable related to MACD. The oscillator signal line (a moving average of the oscillator).

**RSI** : Relative Strength Index

**SfastK** : Stochastic Fast %K (number of periods : 14)

**SfastD** : Stochastic Fast %D (number of periods : 3)

**SslowD** : Stochastic Slow %D (number of periods : 3)

**SMI** : Stochastic Momentum Index

**SMIsignal** : Stochastic Momentum Index signal line

1. **Target Variable**

**Yc** : 해당 시점부터 5 영업일 이후 시점까지의 종가(close) 수익률이 +인 경우는 1, -인 경우에는 0의 값을 가지는 변수로 정의됨.

1. **기타 사항**

거래량이 0인 데이터(공, 휴일)는 제거되었음.

하나 이상의 변수 값이 결측인 자료는 제거되었음.

자료에 날짜 정보는 누락되어 있으나, 이 자료는 시간 순서대로 수집되어 있음.

이를 감안하여,

Training (및 validation)에는 행 인덱스가 1~1700, 2001~3000인 자료를 이용하여 모델을 적합하고,

Testing에는 나머지 자료인 행 인덱스가 1701~2000, 3001~3331인 자료를 이용하여 성능을 평가할 것.

+ 앙상블 최소 1개 사용