

Week 3: 特征工程

<https://github.com/bukosabino/ta>

能够根据open, close, high, low, volume快速构建出股票常用的指数

对股价数据:

1. 将Volume这类极大数值转换到log空间
2. 创建一个新的column用来记录相对前一条数据的涨幅（将数值转换为差值）
3. 创建MA-5数据，并绘制出plot
4. 创建MA-30数据
5. 创建RSI指数
6. 创建5天均方差
7. 创建W%R指数
8. 对每日的收盘价计算出z-score
9. 凭借着你对金融的市场，加入更多你认为有用的指数

参考：论文[Stock Price Forecasting by a Deep Convolutional Generative Adversarial Network](#)中，使用的特征合集

Variable	Description
Close	Closing price of the stock at the end of day
Closed	Difference in the stock price compared to the previous close
Closep	Percentage change in the stock price compared to the previous close
Ope	Opening price of the stock
Oped	Difference in the stock price compared to the previous opening
opep	Percentage change in the stock price compared to the previous opening
high	Highest price reached by the stock on the trading day
highd	Difference in the highest price reached by the stock compared to the previous day
highp	Percentage change in the highest price achieved by the stock compared to the previous day
low	Lowest price reached by the stock on the trading day
lowd	Difference in the lowest price reached by the stock compared to the previous day
lowp	Percentage change in the lowest price reached by the stock compared to the previous day
SMA5	Simple Moving Average of the closing price calculated at 5 days
SMA10	Simple Moving Average of the closing price calculated at 10 days
EMA12	Exponential Moving Average of the closing price calculated at 12 days
EMA26	Exponential Moving Average of the closing price calculated at 26 days
MACD	Moving Average Convergence/Divergence. It is a technical indicator to reveal changes in the strength, direction, momentum, and duration of a trend. It is calculated as the difference between the EMA26 and EMA12
MACDsign	The exceeding of MACD values on MACDsign and vice versa are useful signals to identify possible trend reversals in prices. It is calculated as a 9-period exponential moving average of the MACD line
volume	Number of shares traded on the day
volumed	Difference in the number of traded shares compared to the previous day
volume	Percentage change in the number of traded shares compared to the previous day
RoC13	Rate of Change. The ROC calculates the ratio of today's closing price to the closing price of (in our case) 13 previous days
K15	Stochastic Oscillator. The Stochastic Oscillator compares the closing price of the stock with the range of prices in the considered time period (15 days, in our case)
D5	Simple Moving Average of the values of the variable K15 calculated at 5 days
SMA20C	Simple Moving Average of the closing price calculated at 20 days
BOLlow	Bottom line of the Bollinger Bands
BOLup	Upper line of the Bollinger Bands
BOL	Percentage BOL, a volatility index of the stock. It is constructed as the ratio between the difference of close and BOLlow with the difference between BOLup and BOLlow
MOM12	Momentum, an indicator that measures the change in closing prices. Unlike the ROC, the momentum is calculated as the difference between today's closing price and the closing price of the previous period (12 days, in our case)

通过 yfinance 获取的股价历史数据默认已经进行了向前复权计算，但并不确保其它渠道源的股票价格进行了复权计算。

对生成的特征集进行correlation校验，剔除共线性特征。

