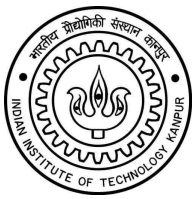

INDIAN INSTITUTE OF TECHNOLOGY KANPUR
DEPARTMENT OF MATHEMATICS & STATISTICS

PROJECT

Generating Trading Signals Using Real-time Time-series Data

Archit Mantri
Bhanu Garg

Supervised by
Prof. Debasis KUNDU

Abstract

Technical indicators are mathematical calculations based on the price, volume, or open interest of a security or contract. By analyzing historical data, technical analysts use indicators to predict future price movements. Examples of common technical indicators include Relative Strength Index, Money Flow Index, Stochastics, MACD and Bollinger Bands. This project focuses on creating technical indicators and then using Machine learning algorithms on the created indicators to predict the stock price.

Signature

April 5, 2019

Contents

1 Introduction

Technical indicators, also known as "technicals", are focused on historical trading data, such as price, volume, and open interest, rather than the fundamentals of a business, like earnings, revenue, or profit margins. Technical indicators are commonly used by active traders, since they're designed to analyze short-term price movements.

This project focuses on using all the information freely available on Yahoo finance about a stock. Using the information about the "Low", "High", "Close", "Open", "Volume" and "Adjusted Close" of a stock, the python code creates 32 technical indicators for the stock. These technical indicators can be used as features in the machine learning algorithms. Idea of the project is to create an automated trading algorithm. Algorithm will have defined set of instructions and it will generate and execute trading signals.

2 Objectives of Study

- To create technical indicators using all the freely available information.
- To apply different machine learning models and predict the stock price.
- To quantitatively analyze the presented strategy and make inference from profit gain using the new strategy.

3 Data Sets

The model is applied on the data sets of the Axis, SBI and Yes Bank. The data set for each bank consists of "Low", "High", "Close", "Open", "Volume" and "Adjusted Close" for each trading day from 2016/09/01 to 2019/01/01. There were total 575 time-series data points for each bank .

The data was obtained from Yahoo Stock Data Source, which is an open repository providing real & static data for various stocks and exchanges.

4 Review of Literature

Smarth Behl[3], in his work uses technical analysis and economic analysis as features for variety of machine learning models to make predictions regarding the stock price movements.

Anthony Macchiarulo [2] in his work, applies technical analysis and machine learning algoritms like SVM, Neural Network and Ensemble method on the 10 years of data of S&P 500 index.

5 Algorithm

The first in the process was to create technical indicators from the freely available data from yahoo finance. A total of 32 indicators were created. To create a balance of every kind, indicators are divided in to 5 categories.

5.1 Momentum Indicators

- **Relative Strength Index (RSI)**
  Compares[1] the magnitude of recent gains and losses over a specified time period to measure speed and change of price movements of a security. It is primarily used to attempt to identify overbought or oversold conditions in the trading of an asset.
- **Money Flow Index (MFI)**
  Uses both price and volume to measure buying and selling pressure. It is positive when the typical price rises (buying pressure) and negative when the typical price declines (selling pressure). A ratio

of positive and negative money flow is then plugged into an RSI formula to create an oscillator that moves between zero and one hundred.

- **True strength index (TSI)**
  Shows both trend direction and overbought/oversold conditions.

- **Ultimate Oscillator (UO)**
  Larry Williams (1976) signal, a momentum oscillator designed to capture momentum across three different timeframes.

- **Stochastic Oscillator (SR)**
  The stochastic oscillator presents the location of the closing price of a stock in relation to the high and low range of the price of a stock over a period of time, typically a 14-day period.

- **Williams %R (WR)**
  Williams %R is a momentum indicator that is the inverse of the Fast Stochastic Oscillator.

- **Awesome Oscillator (AO)**
  The Awesome Oscillator is an indicator used to measure market momentum. AO calculates the difference of a 34 Period and 5 Period Simple Moving Averages. The Simple Moving Averages that are used are not calculated using closing price but rather each bars midpoints. AO is generally used to affirm trends or to anticipate possible reversals.

5.2 Volume Indicators

- **Accumulation/Distribution Index (ADI)**
  The accumulation/distribution line or accumulation/distribution index is a technical analysis indicator intended to relate price and volume in the stock market and acting as a leading indicator of price movements

- **AOn-Balance Volume (OBV)**
  It relates price and volume in the stock market. OBV is based on a cumulative total volume.

- **On-Balance Volume mean (OBV mean)**
  On-balance volume (OBV) is a technical analysis indicator intended to relate price and volume in the stock market. OBV is based on a cumulative total volume.

- **Chaikin Money Flow (CMF)**
  It measures the amount of Money Flow Volume over a specific period.

- **Force Index (FI)**
  It illustrates how strong the actual buying or selling pressure is. High positive values mean there is a strong rising trend, and low values signify a strong downward trend.

- **Ease of Movement (EoM, EMV)**
  It relate an assets price change to its volume and is particularly useful for assessing the strength of a trend.

- **Volume-price Trend (VPT)**
  Is based on a running cumulative volume that adds or substracts a multiple of the percentage change in share price trend and current volume, depending upon the investments upward or downward movements.

- **Negative Volume Index (NVI)**
  The Negative Volume Index (NVI) is a cumulative indicator that uses the change in volume to decide when the smart money is active.

5.3 Volatility Indicators

- **Average True Range (ATR)**
  The indicator provide an indication of the degree of price volatility. Strong moves, in either direction, are often accompanied by large ranges, or large True Ranges.

- **Bollinger Bands (BB)**
  A Bollinger Band is a technical analysis tool defined by a set of lines plotted two standard deviations (positively and negatively) away from a simple moving average (SMA) of the security's price

- **Keltner Channel (KC)**
  Keltner channel is a technical analysis indicator showing a central moving average line plus channel lines at a distance above and below.

- **Donchian Channel (DC)**
  It is formed by taking the highest high and the lowest low of the last n periods. The area between the high and the low is the channel for the period chosen.

5.4 Trend Indicators

- **Moving Average Convergence Divergence (MACD)**
  It is a trend-following momentum indicator that shows the relationship between two moving averages of a securitys price. The MACD is calculated by subtracting the 26-period Exponential Moving Average (EMA) from the 12-period EMA.

- **Average Directional Movement Index (ADX)**
  ADX is used to quantify trend strength. ADX calculations are based on a moving average of price range expansion over a given period of time

- **Vortex Indicator (VI)**
  It consists of two oscillators that capture positive and negative trend movement. A bearish signal triggers when the negative trend indicator crosses above the positive trend indicator or a key level.A bullish signal triggers when the positive trend indicator crosses above the negative trend indicator or a key level.

- **Trix (TRIX)**
  Shows the percent rate of change of a triple exponentially smoothed moving average.

- **Mass Index (MI)**
  It uses the high-low range to identify trend reversals based on range expansions. It identifies range bulges that can foreshadow a reversal of the current trend.

- **Commodity Channel Index (CCI)**
  CCI measures the difference between a securitys price change and its average price change. High positive readings indicate that prices are well above their average, which is a show of strength. Low negative readings indicate that prices are well below their average, which is a show of weakness.

- **Detrended Price Oscillator (DPO)**
  Is an indicator designed to remove trend from price and make it easier to identify cycles.

- **KST Oscillator (KST)**
  It is useful to identify major stock market cycle junctures because its formula is weighed to be more greatly influenced by the longer and more dominant time spans, in order to better reflect the primary swings of stock market cycle.

- **Ichimoku KinkÅ HyÅ (Ichimoku)**
  It identifies the trend and look for potential signals within that trend.

5.5 Other Indicators

- **Daily Return (DR)**
  Calculates the daily return of the stock.
- **Daily Log Return (DLR)**
  Calculates the daily log return of the stock.
- **Cumulative Return (CR)**
  Calculates the cumulative return of the stock.

The next step in the process was to use machine learning models on the created features but 32 features would create a very complex model. Algorithm had two main parameters, F and D. F represent the no of features algorithm will use out of all the created indicators and D represent no of previous days data it will have access to. We randomly chose the Value of F as 7,9 and 11. Value of D was considered 10, 15 and 20. There are 252 trading days in a year and roughly 21 in a month. Because of this, 21 trading days is considered trial period in this algorithm. What it means is for predicting the Value of stock for today, we will train all the 9 models for each of the previous 21 days and the model with the greatest accuracy over this period of 21 days will be used to predict today's stock price.

Using the above algorithm we get an optimum value of F and D for each day and predict the price of stock. We first applied Logistic Regression. Dependent variable was the price of stock converted in binary. If the Stock price was higher today then yesterday's then Dependent variable was given the value 1 otherwise its value was 0. But this model could not differentiate between 1 point change in the stock to say 10 point change. Magnitude of change in the stock was not a dimension.

Next model that we applied was Linear regression. The shortcoming of logistic regression were resolved using this model. Algorithm was applied as it is. Same values of F and D were considered. Dependent variable was the price of the stock. If the algorithm predicted the price of the stock to be greater than today's price then we buy the stock. If the predicted price is less than today's then we sell the stock.

6 Results

We tested the above mentioned Trading Algorithm to conduct trade.We used 510 data points on the test data, which is close to 2 years of trading days.

Yes bank had exceptional return of 69% in 2 years, which is 30% compound annual return.
The success rate of predicting the direction of the stock was 61%.

Axis bank had returns of 35% in 2 years, which is 17% compound annual return.
The success rate of predicting the direction of the stock was 65%.

Sbi had returns of 18% in 2 years, which is close to 8% compound annual return.
The success rate of predicting the direction of the stock was 60%.

References

[1] https://technical-analysis-library-in-python.readthedocs.io/en/latest/.

[2] Anthony Macchiarulo. Predicting and beating the stock market with machine learning and technical analysis, "journal of internet banking and commerce", *CMC Senior Theses. Paper 28*.

[3] Naveed Zaman Smarth Behl, Kiran Tondehal. A machine learning based stock trading framework using technical and economic analysis.