Investment and Trading Capstone Project, Build a Stock Price Indicator

Introduction:

This project relates to investment and trading – namely in the stock, crypto, or commodity markets. Here's where the practice of "buy low and sell high" takes place. However, beyond this simple-minded axiom is a collective complexity creating value for the whole by affording relative market stability and liquidity in ways beneficial to shared economy. It's all easier said then done.

Fundamentally, it is difficult to predict the future and there are an abundant number of opportunities to choose from. Also, there are many different styles of trading and many different types of trading objectives. With the advent of computers, the trading markets are inundated with computational engines that afford trading advantage. A naïve small-time trader who speculates solely on hearsay is likely at a disadvantage against more informed institutional traders whose data center is executing low risk transactions with microsecond response times.

However, in some cases, the future may unfold favorably for the small-time trader. Say the stars align, and the choice to act on opportunity falls congruent to future directions responding favorably. This project is aimed to **promote fun** for such small-time traders in the pursuit of such occurrences.

Beyond such chivalrous reach, this project is an application of Machine Learning principles as an educational exercise. This is not an autonomous trading bot that spits out money. As an indicator, it's just a pointer for market price direction. **Disclaimer:** The application of this project to real world trading is not recommended as results, trading objectives, and trading styles significantly vary.

Project Domain Background:

This project falls into the category of investment and finance. The scope of this project is limited to a small specific context within that domain. Imagine some small fraction of people have some small amount of fun money and they choose to go to Las Vegas and gamble. They win and get excited, or more commonly they lose and walk away humbled; their small amount of money used to gamble with can be lost without causing any major life hiccups. As alternative to going to Vegas, they choose to go on their mobile phone and launch the Robin Hood app. Robin Hood is a mobile phone stock exchange where people can buy and sell stocks for free and so they divert their small amount of fun money into some Robin Hood trade. It's this simple-minded idea to "buy low and sell high", in chance they can get some winning feeling out of this pursuit.

Gaming seems a typical mobile phone thing to do. This sort of speculative gaming seems a small part of investment and finance. To support this gaming fun, is where this project comes in. Say the small-time mobile phone trader first surfs for some price direction information, and then goes to Robin Hood, to better their odds that fun turns into excitement. This project aims to put the odds in favor to the small-time mobile phone trader. Please keep in mind this is not intended as a trading slot machine that demands constant attention. It's more of a day-to-day information update facilitating small short-term trading frames of weeks to months. In this sense, this project falls into the realm of security trading, where a subset of trading seems speculative trading, and where a subset of speculative trading seems mobile gaming. (By way of this gaming slant, a serious tone is softened conducive to education.)

Problem Statement:

Small-time mobile phone-based stock traders face two universal problems: one is that future stock price is difficult to predict, and the second is that there's many trading opportunities to choose from and deciding can be a conundrum. If the future stock price was better understood, and the trading opportunities were ranked (as accessible from a mobile phone), their odds of winning might improve.

Datasets and Inputs:

Historical stock price data is commonly open source. This project does not compare the quality of the free open data sources - and then pick the best one, but simply adopts one that is convenient for educational purposes. (Convenient here suggests compatible with Python programming language upon which the project is built.) In a nutshell, the data is market price versus time. The variables of interest start with the stock ticker symbol, date and market price values such as Open, High, Low, Close, Volume and Adjusted Close. Adjusted Close factors in stock splits and dividend returns. These inputs get pre-processed, then act as input to an algorithmic based Machine Learning Indicator whose output is then fed into a Supervised Machine Learning Classifier. Data can flow in a data pipeline for ongoing update.

Solution Statement:

This project solves the problem of price prediction using a Historical Shape Indicator (HSI), and advises "Buy, Hold, and Sell"; it also strives to rank opportunities via classification to solve the problem of answering the question "which opportunities seem best?"

There are many generic stock price indicator tools available. Relative Strength Index (RSI) is a momentum oscillator indicating overbought and oversold conditions, Moving Average Convergence Divergence (MCAD) indicates trading opportunities via crossings of two Exponential Moving Average (EMA) slow and fast periods. Bollinger Bands afford upper and lower price thresholds that indicate a relative price range tracking the intermediate price trend. These tools can be applied to any market and give standardized perspectives to trading professionals who make trading decisions.

This proposed indicator here differs from these sorts of standardized indicators. This indicator is custom tuned to individual markets. In general, trade returns are considered proportional to the information quotient (known to unknown) plus a square root of the number of opportunities sought. When information quotient goes up, fewer opportunities are needed for the same returns. Likewise, in absence of good information, more opportunities are needed for similar returns. Based on this, the aim of this indicator is to extract information to boost information quotient, boosting trade returns.

From a technical perspective, market price movements contain information as artifacts of trader interaction. Each market has a variety of participants whose resources and objectives vary and as consequence individual markets precipitate a market personality, an embodiment of unique characteristics that distinguish it from other markets having other degrees of unique personality traits, dependent upon their unique market participants. This indicator aims to extract information about market personality by custom tuning (hyperparameters) that capitalize on historical patterns found within the individual market.

"Those who do not learn history are doomed to repeat it." - George Santayana

It seems beyond the reach of most humans to remember market price movement history for each market. This is easier for a computer, where price movement features can be machine learned for increasing information quotient. An algorithmic based Machine Learning Indicator first defines market features, then assesses return probabilities based upon those features. Hence the name Historical Shape Indicator (HSI). Features and probabilities differ depending upon the market personality, hence the approach of tuning the indicator to individual markets. Such a tuned indicator could augment other commonly accessible standardized indicators (RSI, MCAD, BB's). A savvy investor might use an ensemble confluence of indicators to maximize profits. This indicator might play some small part in that, but such an ensemble approach is out of the scope of this project.

An indicator might advise one of three things in a day: Buy, Hold, and Sell. Most small-time mobile investors can only "go long", meaning they buy first then sell later, so the tuned indicator is optimized for this type of behavior. But what if they are tracking a multitude of stock tickers? Here a supervised machine learning approach is pursued to classify opportunities in a way revealing "the best bang for the buck."

Benchmark Model:

The common bench mark model for stock indicator performance is the S&P500 stock index price performance. This is the Standard & Poor's index of the 500 largest US traded companies. The saying "beat the market" is based on the idea of out-performing the S&P500 during similar trading periods. Here, returns during the dates "in the market" are compared, lending indication of benchmark performance to the degree of "beating the market." As the odds go into the favor of the trader, the market gets beat to greater degrees.

Evaluation Metrics:

Beyond bench marking, the Sharpe Ratio is used to ascertain indicator quality, higher as better (but too high as suspicious.) Furthermore, a statistical t-test is used to compare returns from the periods "in the market using the indicator" to randomly selected days "without the indicator". In the case of a strongly trending market, the two may seem similar. However, if machine intelligence lends an edge, the t-test will indicate this as a significant difference in mean returns between these two cases.

Project Design:

Consistent with the recommended approach of Define, Analyze, Implement, Test, and Blog: deliverables include a report and code submissions. With this end in mind, preliminary work includes Use-Case diagrams, Activity Diagrams, a Package diagram, and a Requirements Table to help explain what's going on, and provide "divide and conquer" breakdowns of project components. Following the preliminary scoping, data is scraped, and components are developed using small code block based Jupyter notebooks. An integral part is testing for the attainment of results. Specifically, the algorithmic indicator and ML classifier are scrutinized. Deployment follows, where fruits of a data pipeline can manifest daily in the palm of a mobile-phone based speculative-trade gamer. Finally, requirements are reviewed and discrepancies are reconciled, the final report is completed and the code package is shared for review.

This concludes the proposal; thanks for reading it. 09/22/19 Patrick Parker