Stocks (financial)   Machine Learning   Stock Markets

Can machine learning predict stock prices?

Answer   Request ▾   Follow **204**   Comment   Downvote

16 Answers

William Chen, MS in Applied Mathematics from Harvard
Updated Jul 26, 2016

Yes - many quant trading firms use machine learning techniques on data feeds for automated trades.

These trading firms usually trade on very weak correlations that are uncovered due to research from a quantitative analyst (sometimes known just as a "quant"), who is convinced of the validity of the correlation. While these correlations are weak, the scale at which these quant trading firms operate can make each of these individual strategies worth hundreds of thousands, millions, or even more.

However, there is only room for the fastest firm to make the trade to take advantage of this correlation. Thus, quant firms heavily optimize their strategies for speed. With the low latency, they can beat anyone else to the right trade, since only the fastest players will get the profit.

Since their strategies are optimized for speed and reliability, the machine learning techniques they use are usually very simple. Additionally, because of this ruthless competition for profitable trades, quant firms are incredibly secretive and protective of their intellectual property.

What firms do this?

Some examples of highly-reputable firms that do this include Two Sigma Investments, D. E. Shaw (company), Renaissance Technologies (hedge fund), and Hudson River Trading. These companies are consistently successful in these automated trading strategies, generating very high returns for their clients / themselves.

Consequently, they offer some of the highest compensation packages available in the market (mostly via bonuses) to individuals with the skill-set to identify and execute on these profitable trades.

What's an example of a strategy?

If you want to see evidence of this happening - a Huffington Post blogger found evidence of **Berkshire Hathaway share values increasing whenever Anne Hathaway is mentioned in news.**[1] This is likely an example of some program from a trading firm running automated trades whenever they detect (positive) mentions of "Hathaway" in the news.

While this specific example is humorous since its a false positive, it's a great example of a program running automated trades on a continuously running data feed. The techniques here likely involve data ingestion (they need to read news releases across a

---

Question Stats

203 Public Followers

139,778 Views

Last Asked Sep 9

Edits

negative).

If the algorithm can react to a positive news article faster than anyone else in the market, they can make the profit that is the jump (or decrease) in price.

Can I do this myself?

Probably not. The quant trading firms that do this successfully have spent years perfecting their data sources, infrastructure, talent, risk-management, compliance, raising capital, and all of the other practical necessities to build up a profitable trading operation.

Especially at the extremely quick latencies at which these quant firms discover and execute on signals, it's nearly impossible for any individual (or group of individuals) to do this on their own without dedicating resources to building a company around this. This is an extremely competitive market, from which you won't really find too much help on where to start (which is why the founders of these quant funds tend to have learned from being a part of another quant fund).

See If a Wall Street proprietary trader is able to make lots of money working in a bank, does it mean he can almost surely make money trading on his own account if he quits his job? for more on this.

Read more

This topic in the news: Big Data Gets Bigger: Now Google Trends Can Predict The Market

This topic as a small-scale data science project: What are some actual projects that data scientists have worked on? What tools and analytical techniques were used, and what mistakes were made?

Footnotes

[1] The Hathaway Effect: How Anne Gives Warren Buffett a Rise

Add a comment...   Recommended All

Hristo Piyankov, Data Scientist by heart, manager of Data Scientists by trade
Answered Aug 8

I will go against what everyone else is saying and tell you than **no, it cannot do it reliably.** I have done algorithmic trading and it barely beats an index with a buy and hold strategy or some semi-active trading, as long as you can keep your emotions in check.

As long as everything is going as expected, sure you can have a model with some degree of accuracy, but at the end of the day you do not need an algorithm to tell you that when the trend for the last few months was up, then for the next month it will also be up. You need an algorithm which can reliably predict market corrections and I do not believe that it exists. Reasons to follow:

movements are essentially a random walk. You can look it up, but in a nutshell it means that no one can reliably predict the market, because it's random. The more I am involved into trading, the more I tend to believe this one.

propensity to buy a product, it's most likely that both the company wants to sell the product and customer want's to buy it - we have a consent on both sides. If I am predicting the customer's probability to default, the customer might want to take a loan, but all the decision weather to give it or not, is within the company. In those cases there is no conflict.

3. **Speed**. Related to the point above. If a reliable algorithm for predicting stock movements exists, then you would have no adequate counterpart. When you want to sell, everyone else will want to sell, and vice versa with buying. At the end you will become the market itself and thus no profit would be possible. At the end it would boil down to who can run the algorithm the fastest and I assure you, it will not be an individual user.

4. **Being irrational**. ML algorithms mainly rely on reliable data. Having worked for years, doing modeling in finance, I can tell you people are irrational. Especially when it comes to money. So at the end this data is not reliable.

4.4k Views · 16 Upvotes

Upvote **16**   Downvote

[avatar] Add a comment...   Recommended All

Paul Reuter, Developer. Analyst. Tinkerer. Trader.
Updated Aug 8

I have a recurrent neural net that takes as input the most recent 60 time period outputs from a dozen random forests.

The recurrent neural network was trained with 3 outputs: buy, sell, hold in a reinforcement learning fashion.

The random forests were trained on price-centric features, modified as percentage change of some kind. Each has the same input features, aggregated for a different market sector.

Why do I share this with you? Because it's obvious now that I'm not the only one doing exactly what I'm sharing. My model's performance has deteriorated over time. Market dynamics have changed. Retraining doesn't matter, because there are more players...

Soon, very soon, these predictable signals will yield incredibly swift swings that will blow up a leveraged account. This is a warning.

31.3k Views · 52 Upvotes

Upvote **52**   Downvote

[avatar] Add a comment...   Recommended All

Answered Mar 22, 2015

I have tried to do this using sentiment-analysis and some trend analysis on price
movement. I built a web-crawler in python, and sourced data from the Quandl API
(Find, Use and Share Numerical Data   ), however the results were pretty abysmal
(less than a coin-toss accuracy). I could have used more features or done better data
cleaning, but it became apparent that I was approaching it from the wrong angle.

"Predicting" stock price has ended less-than-favorably for many people, however I
know of several funds that use it in their strategy. Prof. Tucker Balch at Georgia Tech
teaches a class on the subject, and there is some interesting material on his webpage
here: (Tucker Balch's homepage   ).

26.5k Views · 24 Upvotes

Upvote **24**   Downvote

Add a comment...   Recommended   All

Kasper Christensen, studied at Aarhus University
Answered Aug 5

Most data scientist / data analysts have probably wanted to dig into this topic at some
point. This includes me. The reason why is obvious $$$

What I find extremely intriguing about this topic is that I occurred no people who
actually write about what have actually been tried. People are never being specific.
E.g. William Chen speaks as if he knows what the trading firms are doing and he
speaks about trading based on "low correlations". What exactly is meant by "low
correlations"? Does that mean trading by supervised machine learning or what is
meant here? One could also state that traders are trading by low "cosine similarity"
which is essentially the same as a low correlation. My point here is that simply stating
that low correlation and speed is the key is VERY unspecific. If a person have TRUE
insight into how stock prediction is done, the person would (never) write about the
actual methods that are being used and what is the data foundation. Do one need to
have news data? Do one need OTHER stock data than the one, one is trying to predict?
What about social media data? And in addition... Everyone is talking about sentiment
analysis but there are SO many other nuances in the language that negative vs.
positive sentiments. Who looked into that? And what about the methods? How long a
time period do one need to get significant low correlations? What is the best algorithm
for prediction? Support Vector Machines, Neural Networks or maybe Partial Least
Squares?

Obviously if one had a good method one would keep it to oneself. That speaks for the
fact that a good machine learning trading method would never be revealed and one
would have to start from scratch if one want to learn how to do it. For everyone who in
the future want to talk about how machine learning can be used to predict stock
prices, will you please share empirical evidence for your claims.

That being said. My PLS algorithm seems to be doing fairly good on predicting Apples
stock prices on a hold-out test set. So my answer to the question is that it seems
possible based on the data I have. I used R Quantmod, R PLS, and a sliding window
approach. The rest you need to figure out yourself ;)

3.8k Views · 6 Upvotes

Upvote **6**   Downvote

Add a comment...   Recommended   All

I do not believe this is practically possible:

stock price itself is a derivative to the underlying business and its capital structure i.e. both economic and control condition, as well as human emotions and intentions

Mathematically if you take the derivative of some function you are bound to lose some information of the underlying function. Same as stock price, if the only input your algorithm take is the stock price, there are a whole lot information you are going to lose about the underlying factor that will affect the price.

So the only way for machine learning to precisely predict the stock price, you will need to feed ALL the information there is that will affect the stock price, both public and non public. Which is practically impossible to obtain or train a learning algorithm on.

So the answer is no, there is no way to predict stock price using machine learning. You better invest your time and knowledge on something that's more meaningful to mankind than make money in stock market.

329 Views · 4 Upvotes

Upvote   **4**   Downvote

| | Add a comment... | | Recommended   All |

Sam Sachedina, Machine learning and more
Answered Mar 18, 2015

Computational trading is focused more on the speed of strategy execution vs intelligence. Statistical arbitrage, etc.

To answer your question: yes it can. However, you'll make much more if you can execute algorithmic trading strategies FASTER than everyone else.

23.2k Views · 3 Upvotes

Upvote   **3**   Downvote

| | Add a comment... | | Recommended   All |

Milind Paradkar, Senior Content Writer at QuantInsti (2016-present)
Answered Jan 7

Here's link of couple of articles for application of ML in forex markets. Similarly, one can use ML for predicting stock prices.

Machine Learning and Its Application in Forex Markets [WORKING MODEL]

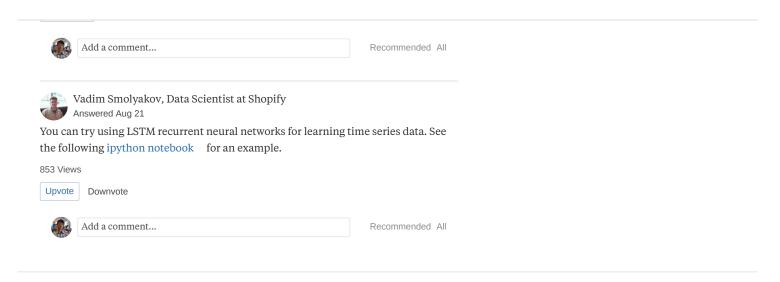**Machine Learning and Its Application in Forex Markets – Part 2 [WORKING MODEL]**

What is required is a good understanding of the process to follow when building a ML model for trading. Then comes the understanding of some of the popular ML algorithms that are used in trading.

Predictive Modeling in R for Algorithmic Trading

One more on Sentiment analysis in trading.

Sentiment Analysis in Trading Using R [WORKING MODEL]

Hope this helps!!

Add a comment...   Recommended All

Vadim Smolyakov, Data Scientist at Shopify
Answered Aug 21

You can try using LSTM recurrent neural networks for learning time series data. See
the following ipython notebook   for an example.

853 Views

Upvote   Downvote

Add a comment...   Recommended All

Top Stories from Your Feed