

Thread from

Replay of Replay of Does machine learning for trading really work with invaluable tips from the likes of London Quant

<http://quantlabs.net/blog/2014/11/replay-of-replay-of-does-machine-learning-for-trading-really-work-with-invaluable-tips-from-the-likes-of-london-quant/>

Here were the questions with the pros vs cons of machine learning !

<http://www.meetup.com/R-Matlab-Users/>

<http://www.meetup.com/quant-finance/>

DOUBTS/CONS

Anyone have any success with Support Vector Machine as in

<http://quantumfinancier.wordpress.com/2010/06/26/support-vector-machine-rsi-system/>

50% success rate in past Meetup presentation

Ernie has doubts

<http://epchan.blogspot.ca/2010/10/data-mining-and-artificial-intelligence.html>

Will,

There are lots of pitfalls in backtesting a strategy even without using machine learning techniques that may render "60% accuracy" quite meaningless. For e.g., are we certain that there is no look-ahead bias in this research? The only way to find out is to replicate the results.

Ernie

Hi Anon,

Thanks for your kind comments.

I have not been able to apply PCA in any profitable pursuit.

"With artificial intelligence, we are summoning the demon," Elon Musk said at the end of last month.

ALso, our Dan Rico was part of this:

<http://www.trade2win.com/boards/attachments/metatrader/130540d1330423251-build-neural-network-indicator-mt4-using-neuroshell-million-model-test.pdf>

Conclusion:

Overall, our research does not corroborate the results reported by some other studies. We find no evidence that SVM techniques enjoy any advantage over traditional methods of forecasting market direction.

Furthermore, none of the common technical indicators tested in our study is able to add predictive power to models based on historical price data alone.

Consistent with the weak form of the Efficient Market Hypothesis, it appears that all relevant information is incorporated in the current price and neither technical indicators nor non-

linear modeling techniques are capable of revealing additional information

, at least for the stocks and indices examined here. An exhaustive test of over 1,000,000 models employing a broad array of non-

linear classification techniques produced excellent results in

sample, but failed to perform well out

-
of
-
sample. This failure
to generalize is a pattern seen in other research and
suggests
that the in
-
sample results are the product
of curve fitting, something that non
-
linear classification methods are often able to
accomplish very well
.
The inference is that the
positive
results reported in some other studies may be the product of
limited
sample selection.

PROS WITH POSSIBILITIES

Supposed hedge funds using machine learning

- Two Sigma (<https://www.twosigma.com>) <-- the dude who caught with a base salary of \$547k
<http://quantlabs.net/blog/2014/11/anoter-quant-analyst-from-mit-accused-of-stealing-trading-model-and-algo-as-a-software-developer-base-pay-of-547k/>
- PDT Partners (<http://www.pdtpartners.com>)
- DE Shaw (<http://www.deshaw.com>)
- DRW Trading Group (<http://drw.com>)

<http://www.quora.com/Is-there-any-hedge-fund-using-Machine-learning-based-algorithms-for-trading>

<http://nips.cc/Conferences/2013/Sponsors/>

How you can apply DotNet Fsharp for machine learning with forecasting trading models

<http://quantlabs.net/blog/2014/11/how-you-can-apply-dotnet-fsharp-for-machine-learning-with-forecasting-trading-models/>

Focus on regresssion, algebrae, stats, etc

Anyone have success with unsupervised

For unsupervised, could you use something like

Matlab System Identification <http://www.mathworks.com/products/sysid/features.html#key-features>

<http://www.mathworks.com/machine-learning/>

Me (to All - Entire Audience):

6:47 PM: hey al

Me (to All - Entire Audience):

6:47 PM: hey all

Christian Laurin (to All - Entire Audience):

6:47 PM: yes

Me (to All - Entire Audience):

6:47 PM: can you hear me ?

ryan (to All - Entire Audience):

6:47 PM: yes

James Hirschorn (to All - Entire Audience):

6:50 PM: I'm studying statistical learning. Not actually "doing" it yet.

James Hirschorn (to All - Entire Audience):

6:50 PM: Haha.

James Hirschorn (to All - Entire Audience):

6:53 PM: Bryan: I have to leave around 8, so I'm hopeful for a recording...

Geoff (to All - Entire Audience):

6:53 PM: Hello everyone from SF

MikeNY (to All - Entire Audience):

6:53 PM: Hi from NY

Daniel Howard (Siesta Key, FL) (to All - Entire Audience):

6:54 PM: Hey Bry

MikeNY (to All - Entire Audience):

6:54 PM: raining

Daniel Howard (Siesta Key, FL) (to All - Entire Audience):

6:54 PM: Cold here. 70 and rain

Daniel Howard (Siesta Key, FL) (to All - Entire Audience):

6:54 PM: brrrrrrrr!

Daniel Howard (Siesta Key, FL) (to All - Entire Audience):

6:55 PM: Paddle boarded today

ryan (to All - Entire Audience):

6:55 PM: where do you post the recordings?

Daniel Howard (Siesta Key, FL) (to All - Entire Audience):

6:55 PM: lol

Me (to All - Entire Audience):

6:55 PM: [youtube.com/quantlabs](https://www.youtube.com/quantlabs)

Me (to All - Entire Audience):

7:00 PM: <http://epchan.blogspot.ca/2010/10/data-mining-and-artificial-intelligence.html>

Sholom (to All - Entire Audience):

7:02 PM: no

Geoff (to All - Entire Audience):

7:03 PM: I used <http://www.nutonian.com/products/eureqa/>, it worked but drawdown was more than strategy made

MikeNY (to All - Entire Audience):

7:04 PM: I dont think that the problem is the machine learning

MikeNY (to All - Entire Audience):

7:04 PM: but the lack of people really known how to use ML

Steven Perrott (to All - Entire Audience):

7:05 PM: most systems don't consistently win either

Christian Laurin (to All - Entire Audience):

7:05 PM: IMO the problem with machine learning is that most forms of it, outside of Watson, assume a non-stochastic process. The market is stochastic IMO...

Steven Perrott (to All - Entire Audience):

7:05 PM: AI is part art and part science

Steven Perrott (to All - Entire Audience):

7:06 PM: it largely depends on the question you put to the system to answer

Steven Perrott (to All - Entire Audience):

7:07 PM: you have to show the system the data needed to answer the question put out

Steven Perrott (to All - Entire Audience):

7:07 PM: yes

MikeNY (to All - Entire Audience):

7:07 PM: yes

Sholom (to All - Entire Audience):

7:07 PM: it can work

Daniel (to All - Entire Audience):

7:07 PM: Yes

Jeff9 (to All - Entire Audience):

7:07 PM: yes (Hi all)

Edwol (to All - Entire Audience):

7:07 PM: y

James Hirschorn (to All - Entire Audience):

7:07 PM: Possible

Ash (to All - Entire Audience):

7:07 PM: yes but the whole system needs to be dynamic

Geoff (to All - Entire Audience):

7:08 PM: Buyins.net uses watson plus other things to picks stocks

Geoff (to All - Entire Audience):

7:08 PM: no

Steven Perrott (to All - Entire Audience):

7:08 PM: easy to make a non working system

Dave (to All - Entire Audience):

7:08 PM: it works 200%

PatrickT (to All - Entire Audience):

7:08 PM: no experience in that field...

Christian Laurin (to All - Entire Audience):

7:09 PM: Geoff... Do you mean IBM Watson?

Christian Laurin (to All - Entire Audience):

7:09 PM: That is what I am referring to

Edwol (to All - Entire Audience):

7:09 PM: y

PatrickT (to All - Entire Audience):

7:09 PM: agree

Geoff (to All - Entire Audience):

7:09 PM: yes

Steven Perrott (to All - Entire Audience):

7:09 PM: even beyond trading

Daniel Howard (Siesta Key, FL) (to All - Entire Audience):

7:09 PM: if human learning works, then the failure is application of ML, as prev stated

Jeff9 (to All - Entire Audience):

7:09 PM: absolutely

PatrickT (to All - Entire Audience):

7:09 PM: agree but necessary for trading :)

PatrickT (to All - Entire Audience):

7:09 PM: not

Me (to All - Entire Audience):

7:10 PM: <http://quantlabs.net/blog/2014/11/how-you-can-apply-dotnet-fsharp-for-machine-learning-with-forecasting-trading-models/>

Steve Eberbach (to All - Entire Audience):

7:11 PM: My opinion of "why machine learning does not work" is "because traders expect the machine to come up with with an answer which always works". The more likely truth is more like the African Watering hole, which ATTRACTS both the winners and the losers, so the algo which works best is like the watering hole without predators. Also, most people think the more data the better the prediction. Too much data, the noise will be learned and swamp the relevant information. So ML must be used very judiciously. As another tool, nowhere near a whole answer. So my answer is yes,BUT...

LQ (to All - Entire Audience):

7:11 PM: ML works but unless you have experience you would end up using 3rd party libraries without understanding wtf you are doing... people cannot blame ml when their system doesn't work

Dave (to All - Entire Audience):

7:11 PM: ML works if your objectives are clear cut

LQ (to All - Entire Audience):

7:12 PM: people should first take a serious class in mathematics before trying to use ML

Daniel (to All - Entire Audience):

7:12 PM: YES!

Christian Laurin (to All - Entire Audience):

7:12 PM: agreed

Mike (to All - Entire Audience):

7:12 PM: LQ, decision trees are just basic programming logic, though, so what's the difference vs. loops and try-catch switches?

Edwol (to All - Entire Audience):

7:12 PM: if you have a good system, use ML to make it better

LQ (to All - Entire Audience):

7:13 PM: decision trees is a general logic

LQ (to All - Entire Audience):

7:13 PM: it comes down to which technique are you using

Steven Perrott (to All - Entire Audience):

7:13 PM: consider the trading system to be an insect....it is taught to gather food....in range it may EAT with BuyLimits and SellLimits while another version feeds by either market orders or BuyStops and SellStops. Set targets above and below and task your system to bias to the long or short side and place trades accordingly.

LQ (to All - Entire Audience):

7:13 PM: ID3, C4.5 C5? hybrid variants?

LQ (to All - Entire Audience):

7:14 PM: there are plenty with different results each

Steven Perrott (to All - Entire Audience):

7:14 PM: They have the finances to give it a strong effort.

LQ (to All - Entire Audience):

7:14 PM: the hedge funds also hire people who are professionals with experience in ML

ryan (to All - Entire Audience):

7:15 PM: if someone uses ml and is successful, people will copy them and then their system won't work anymore

Ash (to All - Entire Audience):

7:15 PM: yess

Jeff9 (to All - Entire Audience):

7:16 PM: Machine learning is defined as the field of study that gives computers the ability without learn without being explicitly programmed.

LQ (to All - Entire Audience):

7:16 PM: I was replying to Mike

Mike (to All - Entire Audience):

7:18 PM: LQ says he doesn't like machine learning blackboxes though, more logical loops like decision trees and genetic programming

Steven Perrott (to All - Entire Audience):

7:18 PM: find out what drives the market in question

Ash (to All - Entire Audience):

7:18 PM: The problem with ml is the same as an other blackbox trading system. the market changes and so what the machine is learning has to change

LQ (to All - Entire Audience):

7:18 PM: the average Joe shouldn't expect to use ML and be successful in trading

LQ (to All - Entire Audience):

7:18 PM: yes Mike

Mike (to All - Entire Audience):

7:18 PM: yeah

Dave (to All - Entire Audience):

7:19 PM: average Joe needs to know what variables are important to their research

Mike (to All - Entire Audience):

7:19 PM: LQ responded to this with yes: LQ says he doesn't like machine learning blackboxes though, more logical loops like decision trees and genetic programming

Dave (to All - Entire Audience):

7:19 PM: and they can use ML to help discover those parameters

LQ (to All - Entire Audience):

7:20 PM: nowadays when most startup firms are referring to ML they are referring PRIMARILY to regression trees, regression forests, decision trees, Bayesian learning etc...

Dave (to All - Entire Audience):

7:20 PM: my tool of choice is genetic programming

LQ (to All - Entire Audience):

7:20 PM: applying ML on timeseries is far trickier

ryan (to All - Entire Audience):

7:21 PM: how would you apply genetic programming to the market?

Steven Perrott (to All - Entire Audience):

7:21 PM: also the time of day and day of week are data....are you at an open, or in the slower times of the day

Mike (to All - Entire Audience):

7:21 PM: LQ: do you make use of technical analysis or linear regression, or do you just feed the system with raw price data?

Jeff9 (to All - Entire Audience):

7:22 PM: Steven Perrott, read LQ's book (Handbook to NeuroEvolution with erlang) to see the future of ML and more broadly AI

LQ (to All - Entire Audience):

7:22 PM: plus most people have no clue on how to clean data

Steven Perrott (to All - Entire Audience):

7:22 PM: thank you

Steven Perrott (to All - Entire Audience):

7:22 PM: train dogs to trade

Steven Perrott (to All - Entire Audience):

7:22 PM: fetch

LQ (to All - Entire Audience):

7:22 PM: cleaning the data is trully an art in order to decrease the dimensionallity of your problem

Ash (to All - Entire Audience):

7:23 PM: Agreed

LQ (to All - Entire Audience):

7:23 PM: no Mike, I don't use any technical analysis or classic linear regression

Dave (to All - Entire Audience):

7:23 PM: Lq: by clean do you mean remove garbage or cherry pick your regimes

Waleed Ayoub (to All - Entire Audience):

7:24 PM: sas

LQ (to All - Entire Audience):

7:24 PM: by cleaning I mean, remove the noise in order to train your model on the core signal

Geoff (to All - Entire Audience):

7:24 PM: python

Christian Laurin (to All - Entire Audience):

7:24 PM: Java,C#

Edwol (to All - Entire Audience):

7:24 PM: python

Daniel Howard (Siesta Key, FL) (to All - Entire Audience):

7:24 PM: C#, C++

Jeff9 (to All - Entire Audience):

7:24 PM: octave/matlab

Ash (to All - Entire Audience):

7:24 PM: matlab/c++

Mike (to All - Entire Audience):

7:24 PM: LQ- for cleaning data, do you use moving averages or linear regression?

LQ (to All - Entire Audience):

7:24 PM: nope

Daniel Howard (Siesta Key, FL) (to All - Entire Audience):

7:24 PM: some think python is the full monty lol

Dave (to All - Entire Audience):

7:25 PM: LQ: can u talk about that. How you liken to isolate the signal.

Christian Laurin (to All - Entire Audience):

7:25 PM: IMO the algorithms are richest and most easy to integrate via Java or C#

LQ (to All - Entire Audience):

7:25 PM: more advanced math techniques

Dave (to All - Entire Audience):

7:25 PM: such as?

Jeff9 (to All - Entire Audience):

7:26 PM: any language that can handle vectorization (parallel programming)

Daniel Howard (Siesta Key, FL) (to All - Entire Audience):

7:26 PM: work in progress here.

LQ (to All - Entire Audience):

7:26 PM: you can use ICA independent component analysis for example, wavelets,... can go on forever here

Steven Perrott (to All - Entire Audience):

7:26 PM: I knew a guy with a system that needed to be run weekly and then it worked throughout the week

Me (to All - Entire Audience):

7:26 PM: Matlab System Identification <http://www.mathworks.com/products/sysid/features.html#key-features>

Mike (to All - Entire Audience):

7:27 PM: I thought wavelets overfit LQ, so you like more PCA?

Christian Laurin (to All - Entire Audience):

7:27 PM: i have in the past

LQ (to All - Entire Audience):

7:27 PM: wavelets do not overfit

Dave (to All - Entire Audience):

7:27 PM: how can you use wavelets when they are non causal

LQ (to All - Entire Audience):

7:27 PM: they simplify / compress the signal

Jeff9 (to All - Entire Audience):

7:28 PM: SVMs are used ofr image recognition (support vector machines) they could possibly be used for pattern recognition to seek out patterns in the market

LQ (to All - Entire Audience):

7:28 PM: no PCA but ICA

Daniel Howard (Siesta Key, FL) (to All - Entire Audience):

7:28 PM: looking at discipulus

LQ (to All - Entire Audience):

7:28 PM: plus there are many ICA variants

Dave (to All - Entire Audience):

7:28 PM: e.g. how do you handle tails on Real time data

Steven Perrott (to All - Entire Audience):

7:28 PM: Better to make the price action fit within the model....rather than work to fit the current price action....in my opinion anyway

LQ (to All - Entire Audience):

7:28 PM: are you asking me Dave?

Dave (to All - Entire Audience):

7:28 PM: Si Señor

Steve Eberbach (to All - Entire Audience):

7:28 PM: A very important thing to decide is what aspect of your objectives do you wish to focus and improve? Are you interested in discovering how useful persistent cycles are for your particular style of trading and instrument you are trading? Do some of several different ones persist enough to be useful? Under what constraints? That example can show why you need to define your problem very specifically before you even start to benefit from the use of ML for that particular aspect.

ryan (to All - Entire Audience):

7:29 PM: can you explain how you would apply genetic programming to the stock market?

LQ (to All - Entire Audience):

7:29 PM: it purely depends on what models you are using

Christian Laurin (to All - Entire Audience):

7:29 PM: The easiest and most flexible would be to use LISP dialect to generate

Me (to All - Entire Audience):

7:29 PM: <http://www.mathworks.com/videos/algorithmic-trading-with-matlab-for-financial-applications-81775.html>

Steven Perrott (to All - Entire Audience):

7:29 PM: you could have a two pronged approach where one AI system works to make profit while the other counters in efforts to manage risk

Christian Laurin (to All - Entire Audience):

7:30 PM: part of the genetic programming problem is that you get hit into dead ends and need to prune

Dave (to All - Entire Audience):

7:30 PM: Asking specifically about wavelet tails changing at their end points

LQ (to All - Entire Audience):

7:30 PM: Bryan I think he is referring to genetic programming Not genetic algorithms

Ad (to All - Entire Audience):

7:30 PM: Is anyone actually running/have a strategy based on ML---what is the performance and what markets/frequency?

Christian Laurin (to All - Entire Audience):

7:30 PM: can I talk about my experience in GP?

Steve Eberbach (to All - Entire Audience):

7:30 PM: A generic question is what aspect of success are you trying to teach the machine to learn? You might start with what "fitness" do you favor if you had the "grail" of genetic programming software?

Dave (to All - Entire Audience):

7:30 PM: eg tail today of wavelets point $N - 1$ might be different at $N-2$

LQ (to All - Entire Audience):

7:31 PM: the wavelet tails depend on what wavelet filter you are using

Dave (to All - Entire Audience):

7:31 PM: do you use a causal wavelet

Christian Laurin (to All - Entire Audience):

7:31 PM: can I talk about genetic programming?

LQ (to All - Entire Audience):

7:31 PM: plus I am using wavelets for HF data

Mike (to All - Entire Audience):

7:31 PM: I know LQ probably likes eurUSD Asian session, which tends to be mean reverting, probably good for your strategy?

Christian Laurin (to All - Entire Audience):

7:31 PM: how do I talk?

Steven Perrott (to All - Entire Audience):

7:33 PM: prune manually?

Dave (to All - Entire Audience):

7:33 PM: u can use parsimony pressure

LQ (to All - Entire Audience):

7:33 PM: Christian you can avoid this problem by hashing the nodes

LQ (to All - Entire Audience):

7:34 PM: but if you hash the nodes you will avoid adding nodes on the first place that may collide

LQ (to All - Entire Audience):

7:35 PM: also it depends on what kind algo you are using for evolutionary programming

LQ (to All - Entire Audience):

7:35 PM: genetic programming, cartesian, gep, mep

LQ (to All - Entire Audience):

7:35 PM: well genetic programming is a tree based

Dave (to All - Entire Audience):

7:35 PM: can I talk

LQ (to All - Entire Audience):

7:36 PM: cartesian, cep and mep are linear

LQ (to All - Entire Audience):

7:36 PM: which avoid these issues to which you are referring

Daniel Howard (Siesta Key, FL) (to All - Entire Audience):

7:36 PM: one method is to increase mutation rate to help min/max problem, then constrain it when global max part of domain is captured

Daniel Howard (Siesta Key, FL) (to All - Entire Audience):

7:38 PM: the genetic version of old school annealing solution

Ash (to All - Entire Audience):

7:38 PM: thats why it has to always change

Christian Laurin (to All - Entire Audience):

7:38 PM: answer is no

Steven Perrott (to All - Entire Audience):

7:39 PM: you could have a spectrum of results tuned for each hour of the day so that the average is now average between 3 and 4 PM

Steven Perrott (to All - Entire Audience):

7:40 PM: have a secondary system to minimize the DD in the transition

LQ (to All - Entire Audience):

7:42 PM: one of the easiest ways, you use decision trees on top of your GP models

LQ (to All - Entire Audience):

7:42 PM: first generate the GP models

LQ (to All - Entire Audience):

7:42 PM: then generate the decision tree to choose the GP model

Mike (to All - Entire Audience):

7:43 PM: LQ talks about ICA/wavelets for isolating signals?

LQ (to All - Entire Audience):

7:43 PM: wavelets to isolate the noise up to a threshold

Christian Laurin (to All - Entire Audience):

7:43 PM: IMO I keep things simple...

LQ (to All - Entire Audience):

7:43 PM: so you don't end up training a model with high vol noise

Mike (to All - Entire Audience):

7:44 PM: no it is the topic dave was talking about for isolating the signal

LQ (to All - Entire Audience):

7:44 PM: volatility

LQ (to All - Entire Audience):

7:44 PM: Dave no, a wavelet will compress your timeseries

LQ (to All - Entire Audience):

7:45 PM: so it doesn't have to go on forever

Steven Perrott (to All - Entire Audience):

7:45 PM: you may be better off using chunky data rather than tick

Steven Perrott (to All - Entire Audience):

7:45 PM: but you can prepare

Ash (to All - Entire Audience):

7:45 PM: yesss!!

LQ (to All - Entire Audience):

7:47 PM: christian it depends the product. Doing market making things can be very different

Steven Perrott (to All - Entire Audience):

7:47 PM: your data is compromised from the start

Ash (to All - Entire Audience):

7:47 PM: yesss

Steven Perrott (to All - Entire Audience):

7:48 PM: the arb may leave tracks

Christian Laurin (to All - Entire Audience):

7:48 PM: exactly

Christian Laurin (to All - Entire Audience):

7:49 PM: 1000% agreed

LQ (to All - Entire Audience):

7:49 PM: I agree

Mike (to All - Entire Audience):

7:49 PM: so a wavelet based trading algorithm is what you're saying LQ, and then you optimize that with decision trees/genetic programming?

LQ (to All - Entire Audience):

7:50 PM: the wavelet simplifies the data

Dave (to All - Entire Audience):

7:50 PM: LQ what wavelet do u use

Jeff9 (to All - Entire Audience):

7:50 PM: so ML only works if you have access to dark pool data or if you can infer darkpool data?

Dave (to All - Entire Audience):

7:50 PM: there are many.

Christian Laurin (to All - Entire Audience):

7:50 PM: no

LQ (to All - Entire Audience):

7:50 PM: Bryan one of the easiest and straight forward way to use ML,...

PatrickT (to All - Entire Audience):

7:51 PM: well, im in that spot that i have algo that profits, but ML to make me avoid DD must really „read the market news, data....to determine regime or upcoming high volatility which will make DD

Steven Perrott (to All - Entire Audience):

7:51 PM: whatever you trade you need to find the key and relevant data that gives some edge

LQ (to All - Entire Audience):

7:51 PM: doing a dumb backtest, get the results and apply the decision tree on the results,... so what you are trying to find is when is your highest chance to lose,.. NOT optimizing parameters

Christian Laurin (to All - Entire Audience):

7:51 PM: that's a given

LQ (to All - Entire Audience):

7:52 PM: so also in this way you are avoiding the parameters recalibration

Christian Laurin (to All - Entire Audience):

7:53 PM: no

Steven Perrott (to All - Entire Audience):

7:53 PM: would be a good use though

Mike (to All - Entire Audience):

7:53 PM: so you just compare the wavelet normalization to the price LQ?

LQ (to All - Entire Audience):

7:53 PM: yeap

LQ (to All - Entire Audience):

7:54 PM: so in a way you are getting a relative volatility of a signal that is not lagged (in comparison to MAs for example)

Steven Perrott (to All - Entire Audience):

7:54 PM: yes, every system starts with a certain viewpoint and becomes a holistic whole in the end.....and may be opposite nearly another good system too

Mike (to All - Entire Audience):

7:54 PM: the mid-price of the top-of-thebook or do you use market depth/12, LQ?

PatrickT (to All - Entire Audience):

7:55 PM: what about ML for news reading, for news trading or avoiding.

Steve Eberbach (to All - Entire Audience):

7:55 PM: Profit is an aspect, so is trend versus reversion/retracing, so is risk management. Each must be addressed with purpose and knowledge of what is relevant to each aspect.

Christian Laurin (to All - Entire Audience):

7:55 PM: yes

LQ (to All - Entire Audience):

7:55 PM: market depth mike

PatrickT (to All - Entire Audience):

7:55 PM: or twitter readings

PatrickT (to All - Entire Audience):

7:56 PM: i just heard

LQ (to All - Entire Audience):

7:56 PM: news reading is only good if you are using a fast news provider

Christian Laurin (to All - Entire Audience):

7:56 PM: they do read news to trade

LQ (to All - Entire Audience):

7:56 PM: and I mean a professional one

LQ (to All - Entire Audience):

7:56 PM: otherwise you always lag

Mike (to All - Entire Audience):

7:56 PM: Do you compare all 20 levels of the book and backtest it?

Dave (to All - Entire Audience):

7:56 PM: LQ. sorry for beating a dead horse. but if I run wavelet on a time series of my entire period compared to a wavelet in a rolling period. how do u handle that.

LQ (to All - Entire Audience):

7:56 PM: bloomberg feed is still slow

Mike (to All - Entire Audience):

7:56 PM: LQ

Steven Perrott (to All - Entire Audience):

7:56 PM: eSignal

Christian Laurin (to All - Entire Audience):

7:56 PM: exagreed lq

LQ (to All - Entire Audience):

7:57 PM: esignal slow as fuck

Mike (to All - Entire Audience):

7:57 PM: ravenpack is one of the cheapest, hehe

Jeff9 (to All - Entire Audience):

7:57 PM: how about reading reading reddit and reddit comments for reading news?

Mike (to All - Entire Audience):

7:57 PM: 10k a month, bloomberg/reuters machine readable feeds (different from the terminals) are 30k a month

Ash (to All - Entire Audience):

7:57 PM: low latency restrictions

PatrickT (to All - Entire Audience):

7:58 PM: agreed

Steve Eberbach (to All - Entire Audience):

7:58 PM: Behavior on news is so different from normal that it may be best to handle it with a specialized strategy. Both at once might be confusing to a machine learning algo.

LQ (to All - Entire Audience):

7:58 PM: just to get good historical equity info from edi it would cost you more than 80k

Dave (to All - Entire Audience):

7:58 PM: Sorry LQ. I meant rolling wavelet. impaired to non rolling wavelet.

Christian Laurin (to All - Entire Audience):

7:58 PM: but reliability is not part of the equation when doing trading on the news

Mike (to All - Entire Audience):

7:58 PM: iqfeed has a one minute delay on news data

Christian Laurin (to All - Entire Audience):

7:59 PM: Quadcore i7 Linux boxes here

LQ (to All - Entire Audience):

7:59 PM: Dave sent me your email and I will get back to you on that because I am having trouble to follow with everything

Geoff (to All - Entire Audience):

7:59 PM: quadcore

Christian Laurin (to All - Entire Audience):

7:59 PM: yes

Geoff (to All - Entire Audience):

7:59 PM: yes

LQ (to All - Entire Audience):

8:00 PM: Meee :)

Steven Perrott (to All - Entire Audience):

8:00 PM: I have heard Xbox type systems due to the graphics computing power

Christian Laurin (to All - Entire Audience):

8:00 PM: If you are doing this sort of work I use things like REDIS to distribute messages

Christian Laurin (to All - Entire Audience):

8:00 PM: Other messaging platforms are slower

Daniel Howard (Siesta Key, FL) (to All - Entire Audience):

8:00 PM: i7. use GPU if speed needed for calc threads

Me (to All - Entire Audience):

8:00 PM: <http://research.microsoft.com/en-us/projects/orleans/>

Mike (to All - Entire Audience):

8:01 PM: are you using all 20 levels of the book, LQ?

LQ (to All - Entire Audience):

8:01 PM: yeap and more Mike

LQ (to All - Entire Audience):

8:01 PM: I am :)

Steve Eberbach (to All - Entire Audience):

8:02 PM: I use many different computers with different software, so I must precalculate long data series then feed the results to another machine to integrate them.

LQ (to All - Entire Audience):

8:02 PM: just use cuda for the heavy stuff as far you can parallelize them

Mike (to All - Entire Audience):

8:02 PM: Do you submit orders at mid-price, how do you simulate the data?

Christian Laurin (to All - Entire Audience):

8:02 PM: interesting lq nickname... flowtrader :)

Me (to All - Entire Audience):

8:03 PM: <http://calvados.di.unipi.it/>

Steven Perrott (to All - Entire Audience):

8:03 PM: multi slice the data and average the settings

LQ (to All - Entire Audience):

8:04 PM: try to optimize the direction, not the rmse

Steve Eberbach (to All - Entire Audience):

8:05 PM: I'm sure there are many more topics to cover! Many of them could take a few years to cover adequately.

Me (to All - Entire Audience):

8:05 PM: i know steve earlbach

Jeff9 (to All - Entire Audience):

8:05 PM: I agree

Mike (to All - Entire Audience):

8:05 PM: How do you submit orders on the book, lq, is it based on mid-price of the top-of-the-book, or is it like a market making multiple level strategy LQ?

Steven Perrott (to All - Entire Audience):

8:05 PM: test, optimize and see how it goes on the remaining data as validation

LQ (to All - Entire Audience):

8:05 PM: passive market making

PatrickT (to All - Entire Audience):

8:05 PM: LQ are you attending any London meetups...like <http://www.meetup.com/Random-Walkers-Quants-In-A-Pub/> ?

Steve Eberbach (to All - Entire Audience):

8:06 PM: To validate, I like to use interleaved sets of sequential "model", optimize, select, and then test.

LQ (to All - Entire Audience):

8:06 PM: nope

PatrickT (to All - Entire Audience):

8:06 PM: pitty

LQ (to All - Entire Audience):

8:06 PM: most of the time when they meet I am usually still at the office

Steven Perrott (to All - Entire Audience):

8:07 PM: I think the system that you would point your AI towards

Christian Laurin (to All - Entire Audience):

8:07 PM: to throw a curve ball into the discussion ;) I don't backtest or forward test. I use statistics and probabilities

PatrickT (to All - Entire Audience):

8:07 PM: thats what u call driven by passion :)

Mike (to All - Entire Audience):

8:07 PM: Does anyone besides LQ use I2 data here?

LQ (to All - Entire Audience):

8:07 PM: I use stress testing and statistical ML

Geoff (to All - Entire Audience):

8:07 PM: no

Christian Laurin (to All - Entire Audience):

8:08 PM: i mean statistical reasoning

Steven Perrott (to All - Entire Audience):

8:08 PM: simpler end of the spectrum

Christian Laurin (to All - Entire Audience):

8:08 PM: not the average stuff

LQ (to All - Entire Audience):

8:08 PM: christian do you imply graph based models?

Christian Laurin (to All - Entire Audience):

8:08 PM: more bayesian type stuff

Geoff (to All - Entire Audience):

8:08 PM: yes

Steven Perrott (to All - Entire Audience):

8:08 PM: yes

LQ (to All - Entire Audience):

8:08 PM: it is still graph.

LQ (to All - Entire Audience):

8:08 PM: good boy :)

Christian Laurin (to All - Entire Audience):

8:09 PM: :)

Steven Perrott (to All - Entire Audience):

8:09 PM: it is important to know what the system is seeing in the debug and optimization of the system

Geoff (to All - Entire Audience):

8:09 PM: yes

Geoff (to All - Entire Audience):

8:10 PM: speed is important

Steven Perrott (to All - Entire Audience):

8:10 PM: just looking at the account equity in relation to the chart and the actions that the system takes may inform you of sweet spots

Steve Eberbach (to All - Entire Audience):

8:10 PM: Humans are good at seeing patterns so a good visual presentation (graph) can be good input to human brain as a bio-machine learner.

Jeff9 (to All - Entire Audience):

8:10 PM: I'm mostly here to learn

Geoff (to All - Entire Audience):

8:11 PM: yes

Dave (to All - Entire Audience):

8:11 PM: complete agree steve

Daniel (to All - Entire Audience):

8:11 PM: YES!

Christian Laurin (to All - Entire Audience):

8:11 PM: yes

Steve Eberbach (to All - Entire Audience):

8:11 PM: This is very useful.

Steven Perrott (to All - Entire Audience):

8:11 PM: my wife gets bored of this stuff

PatrickT (to All - Entire Audience):

8:11 PM: too much info

LQ (to All - Entire Audience):

8:11 PM: OpenCL (not openci)

LQ (to All - Entire Audience):

8:12 PM: OpenCL is nice but they are lacking some important features still

Dave (to All - Entire Audience):

8:12 PM: LQ: u get my email.

LQ (to All - Entire Audience):

8:12 PM: inline assembly

Christian Laurin (to All - Entire Audience):

8:12 PM: lq you do inline assembly?

LQ (to All - Entire Audience):

8:12 PM: yes christian

LQ (to All - Entire Audience):

8:12 PM: quite heavily to be honest

Christian Laurin (to All - Entire Audience):

8:12 PM: lq; hat tip :)

Steve Eberbach (to All - Entire Audience):

8:13 PM: Newer languages have disadvantage of not much development history and libraries.

Geoff (to All - Entire Audience):

8:13 PM: thanks LQ

Dave (to All - Entire Audience):

8:13 PM: thanks LQ

Sholom (to All - Entire Audience):

8:13 PM: Thanks LQ.

Jeff9 (to All - Entire Audience):

8:13 PM: thanks LQ

Christian Laurin (to All - Entire Audience):

8:13 PM: yes thanks lq

Steven Perrott (to All - Entire Audience):

8:13 PM: Thank you for your valuable input.

Steve Eberbach (to All - Entire Audience):

8:13 PM: Thanks! LQ!

PatrickT (to All - Entire Audience):

8:13 PM: yes, thanks LQ ,Dave too

Edwol (to All - Entire Audience):

8:13 PM: thanks Lq

Daniel (to All - Entire Audience):

8:13 PM: Thanx LQ

Daniel Howard (Siesta Key, FL) (to All - Entire Audience):

8:13 PM: Thanks for your great input!!

LQ (to All - Entire Audience):

8:13 PM: you are welcome guys, it is a nice community with cool discussions

Jeff9 (to All - Entire Audience):

8:13 PM: the book is great LQ, still reading it

LQ (to All - Entire Audience):

8:14 PM: give me a sec, trying tp find a book that most people would like

Daniel Howard (Siesta Key, FL) (to All - Entire Audience):

8:14 PM: Love the neuroev book, but doubt will use erlang

Steven Perrott (to All - Entire Audience):

8:14 PM: first you need a system and a viewpoint to set the ML to answer

Geoff (to All - Entire Audience):

8:15 PM: no

Sholom (to All - Entire Audience):

8:15 PM: ML should spit out the PBO as well

Steve Eberbach (to All - Entire Audience):

8:15 PM: I'll second that, Steven.

LQ (to All - Entire Audience):

8:15 PM: just because we were discussing on wavelets, try to check out this book for anyone that wants to learn wavelets the right way

LQ (to All - Entire Audience):

8:15 PM: Ripples in Mathematics by Jensen

LQ (to All - Entire Audience):

8:15 PM: I was trying to find to pass it to Sholom the other day

Daniel Howard (Siesta Key, FL) (to All - Entire Audience):

8:15 PM: Thanks, LQ

Me (to All - Entire Audience):

8:15 PM: <http://www.amazon.com/Ripples-Mathematics-A-Jensen/dp/3540416625>

Sholom (to All - Entire Audience):

8:16 PM: this is for beginners?

LQ (to All - Entire Audience):

8:16 PM: yeap

Steven Perrott (to All - Entire Audience):

8:16 PM: manage risk first, profit second

LQ (to All - Entire Audience):

8:16 PM: just start the book from scratch and you should be able get through everything

LQ (to All - Entire Audience):

8:16 PM: the book also includes lots of code btw

PatrickT (to All - Entire Audience):

8:17 PM: code in ?

Daniel Howard (Siesta Key, FL) (to All - Entire Audience):

8:17 PM: hope not in erlang! lol

LQ (to All - Entire Audience):

8:17 PM: hahahahah

LQ (to All - Entire Audience):

8:17 PM: erlang is nice if you never want to take the system down

Daniel Howard (Siesta Key, FL) (to All - Entire Audience):

8:17 PM: point taken on the robustness for sure! thanks LQ

Christian Laurin (to All - Entire Audience):

8:18 PM: agreed

LQ (to All - Entire Audience):

8:18 PM: I agree with Dave

PatrickT (to All - Entire Audience):

8:18 PM: so im Java...enough?

Daniel Howard (Siesta Key, FL) (to All - Entire Audience):

8:18 PM: fault survival is important in automation

LQ (to All - Entire Audience):

8:18 PM: fine

Steve Eberbach (to All - Entire Audience):

8:18 PM: Whatever language the person teaching you something speaks.

Mike (to All - Entire Audience):

8:19 PM: so you apply different wavelets on all 20 levels of the book LQ?

LQ (to All - Entire Audience):

8:20 PM: no mike

Christian Laurin (to All - Entire Audience):

8:20 PM: depends

LQ (to All - Entire Audience):

8:20 PM: what was the question

Jeff9 (to All - Entire Audience):

8:20 PM: how about datamining areas of the web for news?

Mike (to All - Entire Audience):

8:20 PM: so just a single wavelet on the mid-price of the top of the book?

Daniel Howard (Siesta Key, FL) (to All - Entire Audience):

8:21 PM: narrow band trading on credit spreads for a couple of years

Steven Perrott (to All - Entire Audience):

8:21 PM: there are some option charts that show all the different configurations....condors, spreads etc.....which may give you ideas in non-option trading

LQ (to All - Entire Audience):

8:21 PM: the majority of my work on options was on market making and delta one

Me (to All - Entire Audience):

8:22 PM: <http://www.meetup.com/quant-finance/messages/boards/>

LQ (to All - Entire Audience):

8:22 PM: two wavelets on bid and ask mike

Mike (to All - Entire Audience):

8:23 PM: and how do you handle the rest of the book, just standard deviation off the top of the book??

Geoff (to All - Entire Audience):

8:23 PM: By from SF, thanks everyone

Steve Eberbach (to All - Entire Audience):

8:23 PM: I am just starting to learn about credit spreads; someone wants me to help them with software...

LQ (to All - Entire Audience):

8:23 PM: that comes down to what yo uwant to get out of it Mike

Ash (to All - Entire Audience):

8:23 PM: lq christian. Thanks for your expert inputs

Steve Eberbach (to All - Entire Audience):

8:23 PM: thanks.

LQ (to All - Entire Audience):

8:23 PM: and how large are your positions

Christian Laurin (to All - Entire Audience):

8:24 PM: :)

Daniel Howard (Siesta Key, FL) (to All - Entire Audience):

8:24 PM: Gnight. Thanks Bry, LQ, all

Jeff9 (to All - Entire Audience):

8:24 PM: I have to study wavelets for this week.

Dave (to All - Entire Audience):

8:24 PM: I'd like to learn how to simulate implied vol based on spikes in underlying

Christian Laurin (to All - Entire Audience):

8:25 PM: Modeling binary options

Jeff9 (to All - Entire Audience):

8:25 PM: datamining (for new)

Daniel Howard (Siesta Key, FL) (to All - Entire Audience):

8:25 PM: wavelets to compress input data

Jeff9 (to All - Entire Audience):

8:25 PM: news

Daniel (to All - Entire Audience):

8:25 PM: Genetic programming

LQ (to All - Entire Audience):

8:25 PM: Data cleaning

LQ (to All - Entire Audience):

8:25 PM: :)

Steven Perrott (to All - Entire Audience):

8:25 PM: systems....what is the edge

Dave (to All - Entire Audience):

8:25 PM: Data Cleaning!

Sholom (to All - Entire Audience):

8:26 PM: Graph Models

Steven Perrott (to All - Entire Audience):

8:26 PM: I think the edge comes before you add in ML

LQ (to All - Entire Audience):

8:26 PM: good one Sholom :)

Sholom (to All - Entire Audience):

8:26 PM: :)

Jeff9 (to All - Entire Audience):

8:26 PM: control theory

LQ (to All - Entire Audience):

8:26 PM: ohh nice

Steven Perrott (to All - Entire Audience):

8:26 PM: game theory

Jeff9 (to All - Entire Audience):

8:27 PM: control theory (robotics)

LQ (to All - Entire Audience):

8:27 PM: that my stuff

Dave (to All - Entire Audience):

8:27 PM: information theory

Daniel Howard (Siesta Key, FL) (to All - Entire Audience):

8:27 PM: S domain

LQ (to All - Entire Audience):

8:27 PM: bode plots, Nyse diagrams etc

Steve Eberbach (to All - Entire Audience):

8:27 PM: Yes, game theory. again

LQ (to All - Entire Audience):

8:28 PM: OpenCL is the equivalent of CUDA

Steven Perrott (to All - Entire Audience):

8:28 PM: ML driven risk management

LQ (to All - Entire Audience):

8:28 PM: it is like comparing OpenGL to DirectX

LQ (to All - Entire Audience):Jeff9 (to All - Entire Audience):

8:30 PM: using compound interest?

Steve Eberbach (to All - Entire Audience):

8:30 PM: use logarithmic compounding

8:29 PM: the one is open source the other is more closed

Jeff9 (to All - Entire Audience):

8:31 PM: LQ: due to liquidity issues?

Steve Eberbach (to All - Entire Audience):

8:31 PM: that is the "higher power"

LQ (to All - Entire Audience):

8:31 PM: seen enough idiots trading 50% of the vwap of the day on equities

8:31 PM: that is the "higher power"

LQ (to All - Entire Audience):

8:31 PM: seen enough idiots trading 50% of the vwap of the day on equities

Steve Eberbach (to All - Entire Audience):

8:32 PM: More tips, please!

LQ (to All - Entire Audience):

8:32 PM: yes Jeff

ryan (to All - Entire Audience):

8:32 PM: good place to start learning genetic programming?

Steve Eberbach (to All - Entire Audience):

8:32 PM: here are good tips.

Dave (to All - Entire Audience):

8:33 PM: ECJ

Jeff9 (to All - Entire Audience):

8:33 PM: TA of Stocks & commodities magazine was my first exposure to FX :)

Steven Perrott (to All - Entire Audience):

8:33 PM: how to market a winning system

Steven Perrott (to All - Entire Audience):

8:33 PM: once we get there...t.hen waht

Dave (to All - Entire Audience):

8:33 PM: essentials of metaheuristics

LQ (to All - Entire Audience):

8:33 PM: stephen get it audited and then it is not a problem to find investors