Jim Holmes
Udactiy
Machine Learning
March 15g, 2019

Proposal

Domain Background

Back in 2001 or 2002 I was very serious about becoming a day trader. I visited a few of the proprietary trading firms in Chicago. I even had Toby Crable's day trading book. Over the years I have worked at startups with ex CME and Citadel employees even though we are all just Network engineers.

At first I was looking to research a basket of equities that I could trade based on technical analysis. But was drawn to trading the S&P 500 E-Mini and the Nasdq 100 E-mini futures contracts because of the ability to trade the broader market with single instruments.

I remember when a member named QUAH, on a forum came up with a simple trading strategy called SVS - Something Very Simplistic and S/SVS - Son of Something Very Simplistic from 2002. 17 years ago the only way to test these trading ideas was to paper trade to see if worked. Some people could back test but not many people had the access we have today to test trading strategies.

The VIX Futures Basis: Evidence and Trading Strategies

> A study then demonstrates the profitability of shorting VIX futures contracts when the basis is in contango and buying VIX futures contracts when the basis is in backwardation with the market exposure of these positions hedged with mini-S&P 500 futures positions. The results indicate that these trading strategies are highly profitable and robust to transaction costs, out of sample hedge ratio forecasts and risk management rules. Overall, the analysis supports the view that the VIX futures basis does not accurately reflect the mean-reverting properties of the VIX spot index but rather reflects a risk premium that can be harvested.

Exploiting Term Structure of VIX Futures
Is a machine learning strategy that utilizes the VIX for it's trades. It does use the ES Futures contract as part it's algorithm.

Twenty years ago SVS and SSVS was not a machine learning name.

SomethingVerySimplistic

I'll take a shot at posting a journal of a very simple system that I'm trading. (trading e-mini SP's - should work on NQ's also)

1. Use a minute bar chart
2. Trades will be entered only at specific times - they will be entered at the opening price of the following bars after the open - 3, 5, 8, 13, 21, 34, 55, 89, 144, 233, 377. So, on a normal 9:30 open, trades would take place at 9:33, 9:35, 9:38, 9:43, 9:51, 10:04, 10:25, 10:59, 11:54, 13:23, and 15:47.
3. Use any indicator you like to decide if you will go long or short - doesn't really matter - I use stochastics. Almost anything will work - even something as simple as going in the direction of the previous bar.
4. Once you are filled at the opening price of the bar, enter a limit order to close your position +1.
5. Enter a stop at -1.25. You can reverse at the stop and enter a limit order to close that position at +1.25, leaving you at B/E for that bar's trade (less commish of course).

Son of/SomethingVerySimplistic

The method is basically the same as "SVS" - entry timing based on one chart, direction of entry decided by stochastic in another.

1. The first chart is 272 tick (272T) with a 17,1,17 stochastic on it.
2. The second chart is 2,176 tick (2176T) that is used for timing entries.

-Enter at the open of each 2176T bar, in the direction of the 272T stochastic, %D>%K short, %K>%D, long, just like "SVS".
-Exit limit +1, stop limit -1.25, just like "SVS" - EXCEPT - NO REVERSAL. Do not reverse on the stop, just take the -1.25.
-For ES, there will normally be around 40 or so 2176T bars per day. During the busy times of the day, a new bar is created about every 6 minutes or so - during lunch, they come out about every 13 minutes. Of course, that is all dependent on trade volume.

With recent technological and data processing advances and new machine learning (ML) techniques it would be possible to test these two strategies using modern tools. Even with these technologies getting historical 1 minute futures data is expensive, so I will be using daily close information. Also, I don’t want to “scalp the ES.” I would like to find a longer term trend to hold a position for a specific point gain.

The following is the datasource that was uploaded to github in .csv format.

Quandl SRF-Reference Futures E-mini Source

Quandl Barchart Continuous Futures E-Mini Source

This project intends to apply ML classification techniques to validate a buy or sell signal from these strategies and momentum indicators. I will be using futures data from quandl (barchart).

I will adjust the strategy testing to fit the data i'm able to access - daily close, not tick data. Since futures contracts have an expiration date I will be using the Continuous Futures option and I do have the individual contract dates as well. However, I would like to pursue looking at the individual contracts to see how they perform.

From a personal point of view, I would like to use this as strategy bases for future trading.

Problem Statement

This project's objective is to identify a buy / sell opportunities and position hold times for the S&P 500 E-Mini futures contract based on momentum indicators. Reinforcement learning techniques will be applied - Q Lerner - will be used.

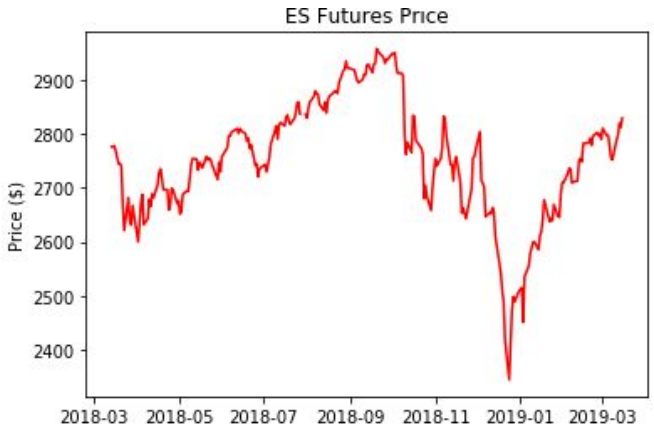
Datasets and Inputs (General view)

Futures market information is by data subscription and is available via api or download to .csv files to backtest, so an educational account will be used to get CME futures data. I will be using data from Quandl (stevens analytics and Barchart). Any comparison to the VIX the data will come from the same source. There is a directory titled futures_data will have all of the .csv files.

Datasets and Inputs (Detailed view)

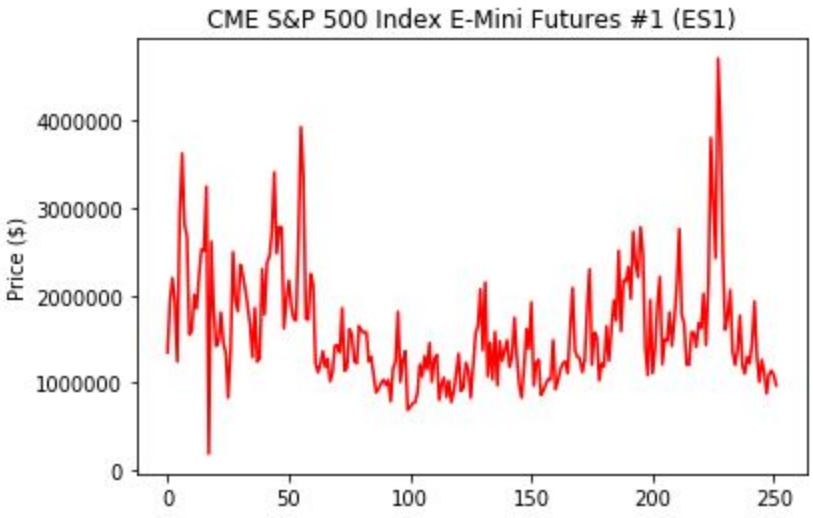
Since I will be looking at Futures data the detailed data source is the same as above. However the data can be broken down into the Futures expiration date calendar or a continuous contract.

Contract expiration chart



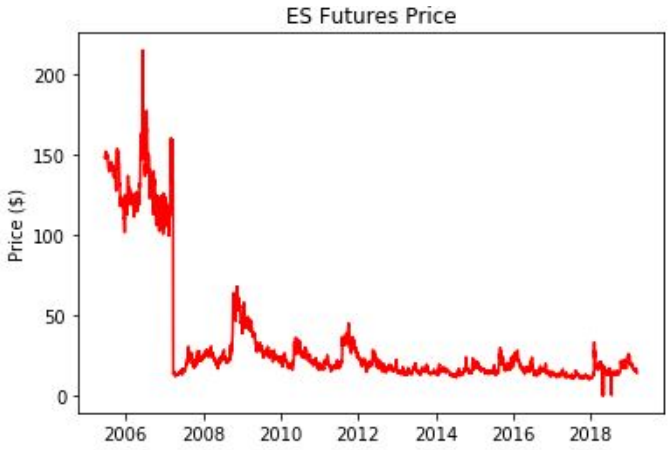
Date	Open	High	Low	Close	Volume	Prev Day OI
2018-03-14	2776.00	2776.00	2776.00	2776.00	0.0	0.0
2018-03-15	2776.50	2776.50	2776.50	2776.50	0.0	0.0
2018-03-16	2777.75	2777.75	2777.75	2777.75	0.0	0.0
2018-03-19	2744.25	2744.25	2744.25	2744.25	0.0	0.0
2018-03-20	2744.50	2744.50	2744.50	2744.50	0.0	0.0
2018-03-21	2739.50	2739.50	2739.50	2739.50	0.0	0.0

Continuous futures



	volume	open	date	\
None				
0	1347509	2500.50	2018-12-31	
1	1943894	2496.50	2018-12-28	
2	2199475	2475.25	2018-12-27	
3	1926447	2353.50	2018-12-26	
4	1239632	2412.75	2018-12-24	
5	2915876	2490.00	2018-12-21	
6	3625398	2510.00	2018-12-20	

Vix continuous



Trade Date	Open	High	Low	Close	Settle	Change	Total Volume
2005-06-20	148.30	148.30	148.30	148.30	148.300	148.30	0.0
2005-06-21	149.00	149.00	149.00	149.00	149.100	0.80	50.0
2005-06-22	148.40	148.40	148.40	148.40	148.600	-0.50	10.0
2005-06-23	149.80	150.50	149.10	150.50	151.200	2.60	61.0
2005-06-24	150.90	152.00	150.50	152.00	152.700	1.50	9.0
2005-06-27	152.80	152.80	152.20	152.20	152.300	-0.40	26.0
2005-06-28	150.20	150.20	150.20	150.20	150.200	-2.10	0.0

Solution Statement

A directional prediction will be developed from the data sources above for a buy or sell signal for a futures contract. Reinforcement learning will be used to identify to buy or sell. A Q Learner will be used to evaluate these metrics.

Benchmark Model

The original trading strategy was tested by paper trading the strategy. This could be done because it was trading a single futures contract, but limited in testing multiple signals to validate a trade. With a machine learning approach we will be able to test different indicators to see which one provides the best buy or sell signal by back testing. This will not be a 1:1 comparison to the original strategy but it will use simple indicators to be back tested.

As a benchmark, a buy and hold strategy for the futures market is monetary suicide - it won't work. A realistic benchmark would be to compare it to a simple supervised learning classification model

Evaluation Metrics

Some of the evaluation metrics will be to use are Discretization and binning so that a Q learner can be applied.

A Sharpe Ratio will be applied as a metric to see how well the strategy performs. Since trading futures has a certain level of risk, It would be interesting to see if a sharpe ratio can help identify trades.

Also, to keep with a simple approach to the strategy an F1 score will be used as well to compare overall performance.

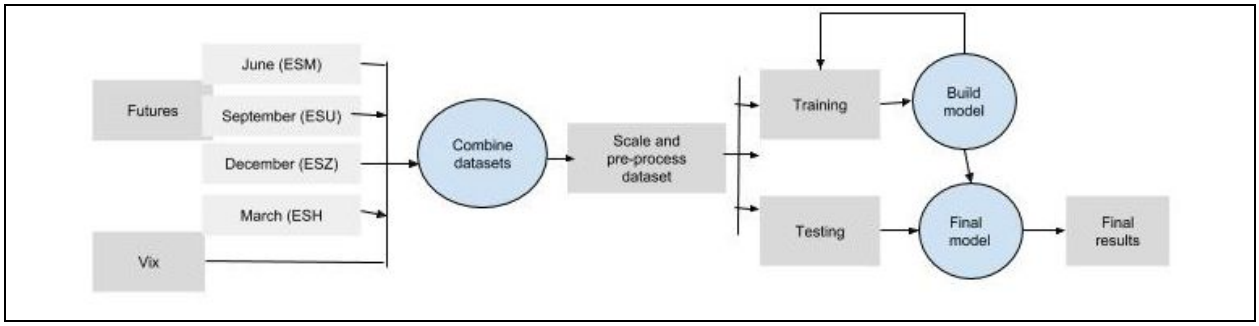
Project Design

A high-level view.

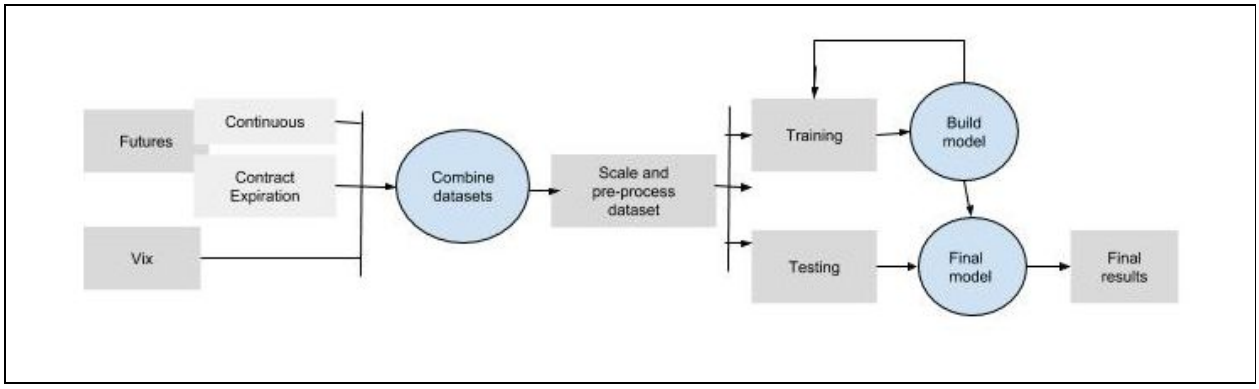
Some of the rewards and penalties could be based on a drawdown. How much of a loss before you close out a position, or how much of a profit is made before you close out a position.

A part of this would be to see if a specific contract over time performs better then others.

- June (ESM)
- September (ESU)
- December (ESZ)
- March (ESH)



Continuous Contract



1. The raw data will be downloaded from Quandl.
2. The data will be pre-processed.
    a. Timeframe will be established.
        i. Open, Close, Volume, Contract.
3. Indicators calculated
    a. Price and Volume
    b. Stochastics
    c. Bollinger bands Percent
    d. VIX applied
4. Apple Discretization and binning
5. Training model with Q learner
6. Test model for to balance the learner speed
7. Compare model for win / loss