

# STOCK OPTION PRICE PREDICTION

ABRAHAM ADAM

## 1. INTRODUCTION

The main motivation for this project is to develop a better stock options price prediction system, that investors as well as speculators can use to maximize their returns.

Many research papers have been written to help investors predict stock price as well as pick portfolio to maximize their returns. However, there aren't much papers that exist for maximizing returns on stock options or predicting stock option prices.

Looking at NIPS and ICML, I saw there is a paper for better coming up with a better stock price predicting, based on the given conditions; however, it assumes that you already know the stock prices for the option.

The goal of this project is to develop a system that will help predict the percent price change for a stock and use that to predict the percent price change of the underlying option for the stock. This can also be used in order to help predict the best option portfolio to help or develop an options trading strategy to maximize returns.

## 2. OPTIONS OVERVIEW

Stock Options are investment instruments that are mostly used in order to minimize risk for a stock trader. An option contract comes in two types:

- (1) 1 Call Options: allows the owner to buy 100 shares of the underlying equity at the price the option specifies.
- (2) 1 Put Options: allows the owner to sell 100 share of the underlying equity at the price the option specifies.

Options have an expiration date, as such the price of an option decreases with each passing day until they become worthless if they are out of the money<sup>1</sup>, or equal to the difference between the price of the equity and option contract if it's in the money<sup>2</sup>. Options come in two flavors - American and European - the difference is that American options can be exercised at any time while European options can only be exercised at expiration. For the purpose of the project we will think of all the options as European options.

---

*Date:* December 14, 2012.

<sup>1</sup>higher than the equity price for call options, and lower than the equity price for put options

<sup>2</sup>lower than the price of the option for call options, higher than the price of the equity for put options

Some speculators and traders choose to use options in order to maximize risk and reward rather than minimize risk. The reason that options as risky is that they are constantly losing value and their price is affected by volatility sense they are intended to counter volatility, if the stock doesn't move the expected direction fast enough, the owner of the option can lose more of the value of the option on time decay and volatility.

### 3. DATA SETS

For the purpose of this project, 2 stocks were chosen as the examples for the general stock market. Those two are:

- (1) Intel which is a DOW 30 component, it represent a large cooperation with low volatility.
- (2) Nvidia Which is a NASDAQ component, it represents a smaller company with high volatility.

For the two stocks chosen, stock volume , open price, closing price, volume, high and lower were obtained from yahoo <sup>3</sup> from the time they went public to December 7th, 2012. This data was used to compute technical indicators that helped us predict stock prices for future days.

Options data were also obtained for both stocks from ivolatility <sup>4</sup> for 6 month of data June 7th, 2012 to December 7th, 2012 - that is 126 trading day, and every day has in the range of 250 different option - different strike price and expiration dates for each call and put. The data that was obtained contains: option type, strike price, mean option price, volume,open interest, delta,vega,gamma,theta,rho<sup>5</sup>.

### 4. STOCK AND OPTION PARAMETERS

**4.1. Stock Indicators.** In order to predict the change in price for a stock, Technical indicators are used to help point the direction we expect the stock to move. below is a small description of the 10 technical indicators that are used.<sup>6</sup>

- (1) Stochastic Oscillator: Location of the current price relative past prices.
- (2) Relative Strength Index: How strong is the stock moving in it's direction.
- (3) Chaikin Money Flow: Measures if money is flowing in or out of the stock.
- (4) Simple Moving Average: Average over the stock prices for a given period.
- (5) Exponential Moving Average: Gives higher value to later values.
- (6) MACD: Difference between two EMA's
- (7) Williams %R: determine if stock overbought over oversold.
- (8) Ultimate Oscillator: combines moving averages from different periods.
- (9) Mass Index: Indicates Reversal based on different from high and low.
- (10) volatility: Standard deviation of a stock over a period.

There is a total of 20 features because different parameters are used in SMA and EMA to capture more of the trend.

<sup>3</sup>nvidia: <http://finance.yahoo.com/q/hp?s=NVDA+Historical+Prices>,  
Intel: <http://finance.yahoo.com/q/hp?s=INTC+Historical+Prices>

<sup>4</sup>[http://www.ivolatility.com/data\\_download](http://www.ivolatility.com/data_download)

<sup>5</sup>explanation of each of those parameters is in the following section

<sup>6</sup>for exact details about how to calculate each of those indicators go to the stockcharts.com link in the references

**4.2. Option Parameters.** When looking at an option it is important to understand how it's price change depending on the underlying equity as well as due to time decay and volatility. Those factors are captured by the greeks for each option.

Delta:  $\frac{\partial \text{Option\_Price}}{\partial \text{Equity\_Price}} =$  change in option priced based on underlying equity

Gamma:  $\frac{\partial^2 \text{Option\_Price}}{\partial^2 \text{Equity\_Price}} =$  sensitivity of delta to underlying equity

Vega:  $\frac{\partial \text{Option\_Price}}{\partial \text{Equity\_Volatility}} =$  change in option priced based on volatility of equity

Theta:  $-\frac{\partial \text{Option\_Price}}{\partial \text{Time\_to\_maturity}} =$  amount lost from option value for every passing day

rho:  $\frac{\partial \text{Option\_Price}}{\partial \text{risk\_free\_interest}} =$  change in option price based on change of riskless interest

## 5. DATA PROCESSING

For stocks, instead of using the stock price as the the expected output for our algorithm, the change in price with respect to the day before is used instead.

The EMA and SMA values that we got are subtracted from the stock price and divided by the stock price so we only have to see the difference of the averages with respect to the stock. That is needed because the expected output is the day's change

The technical indicators for the stock were normalized with respect to the mean and the variance. This was needed because some of the data are stochastic data, so they only oscillate between a certain range, while other are free numbers.

As extra features for each option, the number of days till expiration, and volume of the underlying stock was added. For performance reasons <sup>7</sup> the options parameters set is filtered to contain only the options with strike price within 5 dollars of current price. After filtering we have about 6,000 data points for each stocks.

Black-Scholes <sup>8</sup> was calculated for the options based on the predicted stock prices and it was used in the calculations.

## 6. STEPS FOR STOCK PRICE PREDICTION

The first step for future options price change is being able to predict the price change of the underlying stock, with some reasonable correctness. In order to predict the price change of the future, the technical indicators pointed above are used.

At first an attempt was done to generate regression lines using svm; however, after realizing that the model observed gives over 95% error on the training set, a switch was done to try Support Vector Regression epsilon and nu with Gaussian <sup>9</sup> kernel which provided much better results. The same data sets were with LMS Linear Regression to see if the results are improved on the training set.

<sup>7</sup>running any learning algorithm on the full set takes over 30 minutes

<sup>8</sup><http://en.wikipedia.org/wiki/Black-Scholes>

<sup>9</sup>the function that is chosen is radial basis function

## 7. STEPS FOR OPTION PRICE PREDICTION

Before testing predicted the price of options based on the prices predicted, option prices were predicted based on the true price. That showed high enough results, so a swap was made to the predicted stock price with confidence in the option prediction system.

When predicting the prices three measurements were made in order to see what helps us get the most accurate result. First, measuring options price based on price change and the options parameters listed above. Second, measuring options prices based on Black-Scholes with no learning, just applying the equation. Lastly, measurement was taken for using the result of Black-Scholes as a parameter to SVR and LMS.

For testing the results stated below is based on 80% of the options data going to training set and 20% of the data going to test set.

## 8. RESULTS FOR OPTION PRICE PREDICTION

Correctness is a measurement of the number of predictions that are within a certain margin of error from the true price divided by the total number of predictions. Correctness proves to be a more accurate indicator of how good is the prediction than mean square error because of the large number of options in the test set.

Below C[#] denotes correctness within the #

results with actual stock price ran on the NVIDIA set

Set	MSE	C[0.5]	C[1.0]	C[2.0]
Training Set without Black-Scholes SVR	0.4394	79.7818%	86.4936%	92.7095%
Training Set without Black-Scholes LR	0.3412	70.9960%	74.1602%	77.5303%
Training Set without ML	0.1051	93.5857%	98.4626%	99.8512%
Training Set with Black-Scholes SVR	0.0743	95.5034%	99.2230 %	99.8843%
Training Set with Black-Scholes LR	0.0778	95.8836%	99.3387%	99.8843%
Test Set without Black-Scholes SVR	0.2514	73.3533%	74.6507%	78.2826%
Test Set without Black-Scholes LR	0.0942	59.6806%	78.6427%	93.9122%
Test Set with Black-Scholes SVR	0.02256	96.3562 %	98.7860 %	99.9002 %
Test Set with Black-Scholes LR	0.0576	69.6376%	73.3016 %	78.5502 %

From the results above it's clear that running SVR with with Black-Scholes yields the best result, so when running the learning algorithm on predicted prices for NVIDIA and INTEL we will measure the expected prices using Black-Scholes SVR

Results with predicted prices for NVIDIA and INTC

Set	MSE	C[0.5]	C[1.0]	C[2.0]
Test NVIDIA for next day	0.1011	80.4391%	86.2275%	91.8164%
Test NVIDIA for next 2 day	0.1019	82.7758%	86.6267%	92.2156%
Test NVIDIA for 5 day	0.1203	79.3310%	83.2335%	92.0160%
Test INTEL for next day	0.4389	56.6866%	64.0719 %	75.2495%
Test INTEL for next 2 day	0.1594	54.3912%	62.8743%	74.2515%
Test INTEL for next 5 day	0.1684	53.8922 %	64.2715%	73.5529 %

Results for prediction for NVIDIA's stock price

Set	MSE	C[0.005]	C[0.01]	C[0.02]
Test NVIDIA for next day	0.0247	13.7000%	26.7465%	43.1712%
Test NVIDIA for next 2 day	0.0243	14.4144%	28.0280%	50.9510%
Test NVIDIA for 5 day	0.0243	14.6586%	27.4096%	50.4016 %

## 9. CONCLUSION

With over 50% accuracy we can predict options for next day, 2 days and next 5 days, if 0.5 error is allowed. and nearly 75% accuracy if 2.0 error is allowed. Black-Scholes dramatically improves the correctness of SVR as well as LR, and adding the Greeks adds a slight improvement over just running Black-Scholes with no Machine Learning.

One thing worth noting is that even though the correctness of the predicted option price is low, SVR attempts to make up for the difference and produced better results for the option pricing than we had for the expected stock change. We also notice that increasing the number of days we look in the future doesn't significantly lower the correctness.

In the future, it is worth putting more time into doing a better job predicting the underlying stock prices. The day of the week or month of the year can be added as helpful parameters. analyzing the news may also help yield higher accuracy , as words as "Federal Reserve" or "War" may help us predict the price change better.

## 10. REFERENCES

@articleCC01a, author = Chang, Chih-Chung and Lin, Chih-Jen, title = LIBSVM: A library for support vector machines, journal = ACM Transactions on Intelligent Systems and Technology, volume = 2, issue = 3, year = 2011, pages = 27:1–27:27, note = Software available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm>

@miscwebsite:stockcharts, author = "Interactive Data Corp", title = "Technical Indicators and Overlays", year = "1999-2012", url = "http://stockcharts.com/help/doku.php?id=chart\_school:technical\_indicators"

@miscwebsite:cs229, author = "Ng, Andrew", title = "Machine Learning At Stanford", year = "2012", url = "http://cs229.stanford.edu/materials.html"

@book author = "Haug, Espen Gaardner", title = "The Complete Guide to Option Pricing Formulas", year = "2007", publisher = "McGraw-Hill Professional"