Analyzing The Effects of Bollinger Bands on the Probability of Stock Options Using Support Vector Machines

by

Michael Reeves

A Thesis Presented in Partial Fulfillment of the Requirements for the Degree Master of Science

Approved April 2015 by the Graduate Supervisory Committee:

Andrea Richa, Chair
Daniel McCarville
Hasan Davulcu

ARIZONA STATE UNIVERSITY

May 2015
ABSTRACT

The purpose of this research is to efficiently analyze certain data provided and to see if a useful trend can be observed as a result. This trend can be used to analyze certain probabilities. There are three main pieces of data which are being analyzed in this research: The value for $\delta$ of the call and put option, the %B value of the stock, and the amount of time until expiration of the stock option. The %B value is the most important. The purpose of analyzing the data is to see the relationship between the variables and, given certain values, what is the probability the trade makes money. This result will be used in finding the probability certain trades make money over a period of time.

Since options are so dependent on probability, this research specifically analyzes stock options rather than stocks themselves. Stock options have value like stocks except options are leveraged. The most common model used to calculate the value of an option is the Black-Scholes Model [1]. There are five main variables the Black-Scholes Model uses to calculate the overall value of an option. These variables are $\theta$, $\delta$, $\gamma$, $v$, and $\rho$. The variable, $\theta$ is the rate of change in price of the option due to time decay, $\delta$ is the rate of change of the option’s price due to the stock’s changing value, $\gamma$ is the rate of change of $\delta$, $v$ represents the rate of change of the value of the option in relation to the stock’s volatility, and $\rho$ represents the rate of change in value of the option in relation to the interest rate [2]. In this research, the %B value of the stock is analyzed along with the time until expiration of the option. All options have the same $\delta$. This is due to the fact that all the options analyzed in this experiment are less than two months from expiration and the value of $\delta$ reveals how far in or out of the money an option is.
The machine learning technique used to analyze the data and the probability is support vector machines. Support vector machines analyze data that can be classified in one of two or more groups and attempts to find a pattern in the data to develop a model, which reliably classifies similar, future data into the correct group. This is used to analyze the outcome of stock options. The methodology of how support vector machines work and how they are applied to this study is discussed in chapter 3.
ACKNOWLEDGEMENTS

I wish to express my appreciation to Dr. Andrea Richa for her time and her commitment. I also want to thank Dr. Daniel McCarville and Dr. Hasan Davulcu for their willingness to help in any way necessary. I am grateful to all three committee members for showing interest and participating in this project.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>LIST OF FIGURES</th>
<th>vi</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIST OF TABLES</td>
<td>ix</td>
</tr>
<tr>
<td>LIST OF EQUATIONS</td>
<td>x</td>
</tr>
</tbody>
</table>

## CHAPTER

1 BOLLINGER BANDS AND THEIR EFFECT ON PROBABILITY ……..1

1.1: Introduction……………………………………………………………..1

1.2: Problem Description………………………………………………...2

1.3: The Role of Probability………………………………………………..5

1.4: Our Contribution………………………………………………………..9

1.5: A Strategy For Using Bollinger Bands……………………………11

2 SUPPORT VECTOR MACHINES…………………………………….16

2.1: Methodology: Using Support Vector Machines…………………16

2.2: Linear SVM……………………………………………………………18

2.3: Kernel Methods………………………………………………………27

2.4: Quadratic Programming……………………………………………32

2.5: Previous Uses of Support Vector Machines in Finances……………33

3 TESTS AND ANALYSIS………………………………………………35

3.1: Parameters and Description of Tests……………………………35

3.2: Test Results……………………………………………………………36
<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4 CONCLUSION</td>
<td>71</td>
</tr>
<tr>
<td>4.1: Analysis and Conclusion of Results</td>
<td>71</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>74</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Normal Distribution With Mean of 100:</td>
<td>7</td>
</tr>
<tr>
<td>2. Stock Price Data with Bollinger Bands</td>
<td>13</td>
</tr>
<tr>
<td>3. Linear Support Vector Machine With Linearly Separable Data</td>
<td>19</td>
</tr>
<tr>
<td>4. Linear Support Vector Machine With Non-Linearly Separable Data</td>
<td>19</td>
</tr>
<tr>
<td>5. Support Vector Machine With Polynomial Kernel</td>
<td>28</td>
</tr>
<tr>
<td>7. Parameters</td>
<td>36</td>
</tr>
<tr>
<td>8. S&amp;P 500 2 Month Options Data March 2013 – March 2014</td>
<td>38</td>
</tr>
<tr>
<td>10. S&amp;P 500 April Option Chain April 1 2013</td>
<td>40</td>
</tr>
<tr>
<td>11. SP500 Year SVM Model Polynomial Kernel Degree 2</td>
<td>41</td>
</tr>
<tr>
<td>12. SP500 Year SVM Model Polynomial Kernel Degree 3</td>
<td>42</td>
</tr>
<tr>
<td>13. SP500 Year SVM Model Polynomial Kernel Degree 4</td>
<td>43</td>
</tr>
<tr>
<td>14. SP500 Year SVM Model RBF</td>
<td>44</td>
</tr>
<tr>
<td>15. June 2013 – August 2013 SP500 Data $\delta = 0.2$</td>
<td>45</td>
</tr>
<tr>
<td>16. SP500 Bar Graph $\delta = 0.2$</td>
<td>46</td>
</tr>
<tr>
<td>17. SP500 $\delta = 0.2$ Bar Graph %B Values Greater Than 50</td>
<td>47</td>
</tr>
<tr>
<td>18. SP500 $\delta = 0.2$ Bar Graph %B Values Less Than 50</td>
<td>48</td>
</tr>
<tr>
<td>19. SP500 $\delta = 0.2$ Pie Graph</td>
<td>49</td>
</tr>
<tr>
<td>20. SP500 $\delta = 0.2$ Pie Graph %B Values Greater Than 50</td>
<td>50</td>
</tr>
<tr>
<td>21. SP500 $\delta = 0.2$ Pie Graph %B Values Less Than 50</td>
<td>50</td>
</tr>
</tbody>
</table>
Figure

22. Calls SP500 $\delta = 0.2$ Data Polynomial Kernel Degree 2. ..............................................52
23. Calls SP500 $\delta = 0.2$ Data Polynomial Kernel Degree 3. ..............................................52
24. Calls SP500 $\delta = 0.2$ Data Polynomial Kernel Degree 4. ..............................................53
25. Calls SP500 $\delta = 0.2$ Data RBF Kernel .................................................................53
26. Calls June 2013 – August 2013 SP500 Data $\delta = 0.3$ ..............................................55
27. Calls SP500 $\delta = 0.3$ Data Polynomial Kernel Degree 2. ..............................................55
28. Calls SP500 $\delta = 0.3$ Data Polynomial Kernel Degree 3. ..............................................56
29. Calls SP500 $\delta = 0.3$ Data Polynomial Kernel Degree 4. ..............................................56
30. Calls SP500 $\delta = 0.3$ Data RBF Kernel .................................................................57
31. Puts June 2013 – August 2013 SP500 Data $\delta = 0.2$ ..............................................57
32. Puts SP500 $\delta = 0.2$ Data Polynomial Kernel Degree 2. ..............................................58
33. Puts SP500 $\delta = 0.2$ Data Polynomial Kernel Degree 3. ..............................................58
34. Puts SP500 $\delta = 0.2$ Data Polynomial Kernel Degree 4. ..............................................59
35. Puts SP500 $\delta = 0.2$ Data RBF Kernel .................................................................59
36. Puts June 2013 – August 2013 SP500 Data $\delta = 0.3$ ..............................................60
37. Puts SP500 $\delta = 0.3$ Data Polynomial Kernel Degree 2. ..............................................60
38. Puts SP500 $\delta = 0.3$ Data Polynomial Kernel Degree 3. ..............................................61
39. Puts SP500 $\delta = 0.3$ Data Polynomial Kernel Degree 4. ..............................................61
40. Puts SP500 $\delta = 0.3$ Data RBF Kernel .................................................................62
41. Puts September 2013 – November 2013 BTU Data $\delta = 0.2$ ......................................65
42. Puts BTU $\delta = 0.2$ Data Polynomial Kernel Degree 2. ..............................................65
43. Puts BTU $\delta = 0.2$ Data Polynomial Kernel Degree 3. ..............................................66
Figure

44. Puts BTU $\delta = 0.2$ Data Polynomial Kernel Degree 4............................................66
45. Puts BTU $\delta = 0.2$ Data RBF Kernel.................................................................67
46. Puts September 2013 – November 2013 BTU Data $\delta = 0.3$.................................67
47. Puts BTU $\delta = 0.3$ Data Polynomial Kernel Degree 2.............................................68
48. Puts BTU $\delta = 0.3$ Data Polynomial Kernel Degree 3.............................................68
49. Puts BTU $\delta = 0.3$ Data Polynomial Kernel Degree 4.............................................69
50. Puts BTU $\delta = 0.3$ Data RBF Kernel.................................................................69
## LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Bullish Market Test Results</td>
<td>63</td>
</tr>
<tr>
<td>2. Bearish Market Test Results</td>
<td>70</td>
</tr>
<tr>
<td>Equation</td>
<td>Page</td>
</tr>
<tr>
<td>----------</td>
<td>------</td>
</tr>
<tr>
<td>1. %B Value Equation</td>
<td>4</td>
</tr>
<tr>
<td>2. Linear SVM Equation Group 1</td>
<td>21</td>
</tr>
<tr>
<td>3. Linear SVM Equation Group 2</td>
<td>21</td>
</tr>
<tr>
<td>4. Linear SVM Equation</td>
<td>21</td>
</tr>
<tr>
<td>5. Optimization Problem</td>
<td>21</td>
</tr>
<tr>
<td>6. Lagrange Multiplier</td>
<td>22</td>
</tr>
<tr>
<td>7. Lagrange for Optimal ( w )</td>
<td>22</td>
</tr>
<tr>
<td>8. Lagrange for Optimal ( w ) Solved</td>
<td>22</td>
</tr>
<tr>
<td>9. Lagrange for Optimal ( b )</td>
<td>22</td>
</tr>
<tr>
<td>10. Optimal Lagrange</td>
<td>23</td>
</tr>
<tr>
<td>11. Simplified Optimal Lagrange</td>
<td>23</td>
</tr>
<tr>
<td>12. Constraint I (( \alpha ) Constraint)</td>
<td>23</td>
</tr>
<tr>
<td>13. Constraint II</td>
<td>23</td>
</tr>
<tr>
<td>14. Linear SVM Equation Group 1 With Slack Variables</td>
<td>25</td>
</tr>
<tr>
<td>15. Linear SVM Equation Group 2 With Slack Variables</td>
<td>25</td>
</tr>
<tr>
<td>16. Linear SVM Equation With Slack Variables</td>
<td>25</td>
</tr>
<tr>
<td>17. Optimization Problem With Slack Variables</td>
<td>25</td>
</tr>
<tr>
<td>18. Lagrange Multiplier With Slack Variables</td>
<td>25</td>
</tr>
<tr>
<td>19. Optimal Value of ( C )</td>
<td>26</td>
</tr>
<tr>
<td>20. Optimal Value of ( C ) Solved</td>
<td>26</td>
</tr>
<tr>
<td>21. Optimal Lagrange With Slack Variables</td>
<td>26</td>
</tr>
<tr>
<td>Equation</td>
<td>Page</td>
</tr>
<tr>
<td>-------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>22. Simplified Optimal Lagrange</td>
<td>26</td>
</tr>
<tr>
<td>23. Constraint 1 With Slack Variables</td>
<td>26</td>
</tr>
<tr>
<td>24. Constraint 2 With Slack Variables</td>
<td>26</td>
</tr>
<tr>
<td>25. Constraint 3 With Slack Variables</td>
<td>26</td>
</tr>
<tr>
<td>26. Kernel Dot Product</td>
<td>30</td>
</tr>
<tr>
<td>27. Polynomial Kernel</td>
<td>30</td>
</tr>
<tr>
<td>28. Radial Basis Function Kernel</td>
<td>30</td>
</tr>
<tr>
<td>29. Sigmoid Kernel</td>
<td>30</td>
</tr>
<tr>
<td>30. Kernel SVM Equation Group 1</td>
<td>30</td>
</tr>
<tr>
<td>31. Kernel SVM Equation Group 2</td>
<td>30</td>
</tr>
<tr>
<td>32. Kernel SVM Equation</td>
<td>30</td>
</tr>
<tr>
<td>33. Lagrange for Optimal $w$ Using Kernel</td>
<td>31</td>
</tr>
<tr>
<td>34. Optimal Lagrange Using Kernel</td>
<td>31</td>
</tr>
<tr>
<td>35. Simplified Optimal Lagrange Using Kernel</td>
<td>31</td>
</tr>
<tr>
<td>36. Constraints on Kernel</td>
<td>31</td>
</tr>
<tr>
<td>37. Kernel SVM Equation With Slack Variables Group One</td>
<td>31</td>
</tr>
<tr>
<td>38. Kernel SVM Equation With Slack Variables Group Two</td>
<td>31</td>
</tr>
<tr>
<td>39. Kernel SVM Equation With Slack Variables</td>
<td>31</td>
</tr>
<tr>
<td>40. Optimal Lagrange Using Kernel with Slack Variables</td>
<td>31</td>
</tr>
<tr>
<td>41. Simplified Optimal Lagrange Using Kernel</td>
<td>31</td>
</tr>
<tr>
<td>42. Constraints on Kernel With Slack Variables</td>
<td>32</td>
</tr>
</tbody>
</table>
Chapter 1

Bollinger Bands and Their Effects On Probability

1.1: Introduction

When it comes to the market, one of the common fallacies is traders try to predict the market. Traders try to guess when a stock or an index will increase or decrease in value. How can a rich investor look at a stock and accurately predict when to go in and out of the market? These questions are the wrong questions to be asking.

Many believe that the professional investor is able to accurately predict the future. To the professional investor, this is not the case. The professional investor does not try to predict, he or she analyzes probabilities. Analyzing probabilities does not tell an investor the outcome of a stock each day the market opens. Probabilities basically tell an investor how stocks or stock options perform if they were tested enough times. In other words, if the calculated probability is 0.52 that a stock will perform in a certain way, then if that stock is tested enough times, roughly 52% of the time the stock should perform that particular way. Probabilities do not give information for each individual period; rather, they give an investor the chance a stock should perform in a certain way. If the probability is accurate or close to being accurate, then if the stock is tested enough times, the probability shows how often the stock should perform in the way it was tested. By doing this, an investor can make money rather than worrying about predicting a stock accurately every time. To the investor, what is important is making money in the long run, not predicting stocks correctly. The investor tries to put him or herself in the position of the casino owner who makes money rather than the casino visitor who is concerned about winning or getting lucky. Sheldon Natenberg, a professional investor in stock
options states, “In essence, then, all trading decisions are based on the laws of probability” [1].

The question is how do we analyze the probabilities? There are many effects an investor should consider due to the uncountable forces that drive the market. Different strategies work for different people. But the one goal professional investors have in common is their goal of making money not predicting all stocks correctly. This paper presents and analyzes a strategy that can be used when analyzing stock options.

1.2: Problem Description

Like stocks, stock options have value. Each option on the market involves two people. One investor owns the right to the option and the other has an obligation. The obligator sells an option to a person who is willing to buy a right to buy or sell a stock at a particular price. For example, if Apple is currently selling stock at $100 per share and someone buys a call option of the stock at a strike price of $120 per share, that investor can buy Apple’s stock at $120 per share no matter what Apple’s stock value becomes. This means if Apple’s stock increases in value and its worth becomes $130 per share, the investor who bought the call option at the $120 strike price, can still buy the stock at $120 per share. The buyer buys the option from the obligator who is obligated to sell the stock to the investor at the strike price if the buyer chooses to exercise the option. However, the investor who bought the right would only exercise that option if Apple’s stock reaches a price above $120 per share. This is an example of a call option, which is an option where someone buys the right to buy a stock at a particular price when they believe the stock value will increase. A put option is simply an option where a person
buys the right to sell a stock at a particular price. Stock options are not always exercised. As mentioned before, stock options themselves have value and can be traded like stocks and their value is dependent on the probability the option will be exercised according to the Black-Scholes model. All options have an expiration date, which is the date the option must be exercised by. If the option is not exercised by that particular date, it becomes worthless and can no longer be exercised [2].

If an investor buys a call option which grants him or her the right to buy that stock at a particular strike price and the price of the corresponding stock increases to above that strike price before the expiration date, the investor can exercise his or her right and buy the stock at the strike price even when the actual price of the stock is higher. When this happens, the option expires in-the-money. If an option is not in-the-money, then it is out-of-the-money. For put options, being in-the-money is having the stock price be worth less than the strike price and out-of-the-money when the stock price is above the strike price. Basically, in-the-money for both call and put options means the person who owns the right to buy (in the case of a call option) or sell (in the case of a put option) a stock at a particular price will make money from selling or buying the stock if he or she chooses to exercise that option. The obligator, who originally sold the option, makes money if the option remains out-of-the-money. However, once the option is in-the-money, the buyer of the option begins making money. The amount he or she makes depends on how far in-the-money the option moves [3].

Each option contains values for certain variables to help show the overall value of the option. One of these variables is $\delta$, which represents the change in value of the option for every dollar the corresponding stock increases in value. The value, $\delta$, can be used to
analyze probability due to the fact that \( \delta \) increases the more it goes in-the-money [4]. When an option is in-the-money, the value \( \delta \) is over 0.5, whereas if the option is out-of-the-money, \( \delta \) is below 0.5.

The purpose of this research is to use the value \( \delta \) of a stock option as a reference to see how far in or out-of-the-money an option is and to see if the price of a stock in relation to its Bollinger bands affects the actual probability an option will expire in the money.

The \%B value is used in this study and is a value which shows where the current price of a stock is in relation to the Bollinger bands [5]. Bollinger bands were developed by John Bollinger in the 1980’s. Many stock traders use this indicator when buying and selling stocks. Bollinger bands consist of three lines: an upper band, a lower band, and a middle band. The upper and lower bands are two standard deviations out of the simple moving average over the last twenty days. The middle band is the simple moving average over the last twenty days [6]. If the current price of a stock is at the upper band, the \%B value equals 100%. If the current price is at the lower band, the \%B value is 0%. If the current price of a stock is at the simple moving average, the \%B value is 50%. It is possible for the \%B value to be above 100% if the current price of the stock is above the upper band and be negative or below 0% if the current price is below the lower band. The \%B value is calculated by the following formula: [7]

\[
\%B = \frac{\text{current stock price} - \text{lower band price}}{\text{upper band price} - \text{lower band price}}
\]  

(1)
One strategy that investors use for stocks is to use the Bollinger bands to indicate whether the price of the stock is over or undervalued. However, it is important to recognize when a stock can be recognized as over or undervalued. For example, in a bullish market, stocks tend to not be over valued very often due to the fact the trend for the stock is going up, but it does tend to be undervalued during a time of retracement. The opposite tends to happen in a bearish market where stocks tend to not be undervalued very often due to the fact the trend for the stock is going down, but it can be overvalued during a time of retracement [6]. For this study, this particular strategy is observed and studied. If the market is bullish, we observe low %B values and observe and see if this has an affect on the probability a call or put option will expire in or out of the money. If the market is bearish, we observe high %B values and observe and see if this has an affect on the probability a call or put option will expire in or out of the money.

1.3: The Role of Probability

What has already been mentioned is how important probability is when looking at the market and how investors only care about probability not about winning or losing. One of the greatest investigators of probabilities dealing with stocks and stock options is Sheldon Natenberg. Mr. Natenberg started a career in trading when he became an independent market maker in the Chicago Board Options Exchange. He began trading at the Chicago Board Trade and then became Director of Education for Chicago Trading Company. Mr. Natenberg makes a point that there is a difference between understanding the complete theoretical versus understanding the arbitrage that affects the actual probability. One of the reasons why Mr. Natenberg likes dealing with stock options is
theoretically the probability of the market going up is just as likely as the probability the market will go down. This means that if someone buys a stock, assuming the arbitrage of the market was to be completely disregarded, after making many trades, no money would be made since all the money gained from the market going up would be lost from the market going down just as often. This of course does not include the closing costs of buying the stocks. The point is there is a 50% chance that in two months the stock price then will be higher than the current stock price, but there is also a 50% chance the stock price will be lower than the current stock price. If an investor continuously traded stocks, theoretically, he or she would not gain or lose any money if closing costs are not included.

Of course people in the real world still make money in the market (as well as lose money in the market) rather than breaking even. What causes this to happen? Of course the reason is because the theoretical does not take into account the arbitrage that affects the market every day. Arbitrage means outside real-world effects that influence a stock’s price. Some examples of this are the earning reports of companies in the market, whether certain countries in the world are at war, sudden increase or decrease in demand or supply of what the company sells, etc. All these influences plus other outside effects cause people to see how the price of a stock will be affected in the future. Nevertheless, the point is that initially, before looking at the arbitrage, buying or shorting stocks has no advantage or disadvantage.

With stock options, this is not the case, especially in selling options. To illustrate this point, observe Figure 1, which assumes a normal distribution curve to model the probability of all the possible values of the Apple stock mentioned in the previous
section. As a reminder, assume Apple’s current stock is worth $100 a share. If an investor bought the stock and sold it in two months, theoretically, there is a 50% chance the stock will be worth more than $100 a share and a 50% chance the stock will be worth less than $100 a share. The further away the possible future price is from the current price, the lower the probability that will be the ending price of the stock. For example, there is a greater chance the future price will be worth between $100 and $110 a share than being worth between $110 and $120 a share. Similarly for prices lower than $100, there is a greater chance the future price will be worth between $90 and $100 a share than being worth between $80 and $90 a share. The current price is the mean with the probabilities of future prices continuing to decline as the corresponding price becomes further away from $100 a share as shown in Figure 1 below.

![Figure 1](image.png)

What is highlighted in blue shows the range of possible values within one standard deviation of the mean or current price of the stock. For this example, to make it easy, assume the standard deviation is plus or minus 20 as shown in the figure. This
means there is a 68% chance the price of the stock in two months will be in between $80 and $120 a share assuming the model above is correct.

    With this in mind, rather than buying Apple’s stock, suppose a seller decides to sell a contract of 100 shares at a strike price of $120 a share and a buyer is willing to buy the contract. This would mean the seller is selling a call option where he or she makes money immediately from selling the option. The seller has put himself under an obligation with the contract sold where he would have to sell 100 shares of Apple’s stock to the buyer at 120$ a share if the buyer chooses to exercise the contract. This would only be done if Apple’s stock moves above $120 a share. If the stock does not reach that price by the expiration date of the contract, then the seller has just made his money from selling the contract and the buyer never exercises the contract. However, if the price does pass $120 a share, then the buyer begins making money. Once the price has reached a point where the buyer is now making more money from the value of the option than what he spent to buy the option in the first place, he is now making money.

    Assuming Figure 1 is accurately portraying the probability of Apple’s stock price by the expiration date of the option, the seller would make money 84% of the time. This is the case because the probability the resulting price is any value within one standard deviation is 68%, which would make the remaining probability 32%. This means there is a 16% chance the stock will be below $80 a share and a 16% chance the stock will be above $120 a share. If the stock goes down below $80 a share, the seller still collects his money. The 16% chance the stock will be above $120 a share is the only time the buyer would begin to make money. The seller in this case is playing off of probability whereas the buyer is hoping he or she will be lucky and correctly predict the future price of the
stock. The buyer could be right, but the probabilities show that if many trades similar to this are done continuously, the seller will make money more often than the buyer assuming no arbitrage.

Of course the arbitrage in the real world affects the theoretical probability, but the fact is initially an investor can put himself or herself in a position where the probabilities are in his or her favor. If the stock was simply bought or shorted instead of buying or selling options, the investor has put himself in an initial position of having a 50% chance of making money and a 50% chance of losing money. However, if he sold an option’s contract instead, he has increased the odds from 50% to 84%. The theoretical probability is symmetrical in the case of buying the stock but it is asymmetrical in favor of the investor in the case of selling an option. Of course the arbitrage and other factors do affect these numbers, however, the underlying point is that it is much easier to structure the probabilities in the case of stock options rather than the case of buying or shorting stocks themselves [1].

1.4: Our Contribution

Support vector machines are used in this study to analyze the data. To be able to analyze the data efficiently, a computer program has been written for this study. The program observes data regarding stocks and observes a document that the user inputs in the program which contains the following information: the closing prices of the stock being studied, the dates of those closing prices, and the strike price of each option being studied. The user directly inputs to the program whether the options in the study are call or put options, the date of expiration of all the options in the study, and the closing price
of the stock on the day of expiration. All the options in a particular study have the same value \( \delta \), which means there are different options with different strike prices that have a value \( \delta \) on different days. As the market changes, one option with a particular strike price has a particular \( \delta \) on one day but has a different \( \delta \) on another day. The program analyzes all this data to output certain information. The stock prices are analyzed to calculate the Bollinger bands. By having the Bollinger bands plus the twenty day moving average, the \( \%B \) value for any particular day can be calculated. The program calculates the twenty-day moving average for each day and uses that to find the price that would be two standard deviations above the moving average and two standard deviations below the moving average. The dates of the closing prices are observed to calculate the number of days until the option for that day expires. The strike prices are observed to be able to know whether that particular option expired in or out-of-the-money by referencing the price of the stock on expiration.

The data is placed onto an \( x \)-\( y \) axis. The \( \%B \) is placed on the \( y \) axis and the days to expiration are placed on the \( x \) axis. Every option that expires out-of-the-money is colored blue and every option that expires in-the-money is colored red. The support vector machines observe the data and which options expired out-of-the-money and which options expired in-the-money. They then develop a model based off the data to predict future options based off the \( \%B \) values and the number of days until expiration. The predictions are then analyzed and compared with the actual results to see the accuracy of the model. Support vector machines are covered in more detail in the next chapter.
1.5: A Strategy For Using Bollinger bands

As mentioned before, the main focus of this study is looking at the effects of Bollinger bands. The support vector machines are the tools being used in this study to measure the effect of the Bollinger bands. When using SVM’s in this study, Bollinger bands are one of the two dimensions in the SVM’s in this study.

In both bullish and bearish markets, retracements tend to happen with stocks. A retracement is a temporary pullback for a stock. For example, in a bullish market, a retracement would be the time period when the stock is falling after it has had a long period of increasing in price. In a bearish market, a retracement would be when the stock price is temporarily increasing after having a long period of decline. Retracements tend to happen for only a short period of time. During a bullish market, a stock might temporarily go down after a long period of going up and then continue to go up once the retracement concludes. Investors who use this strategy try to buy when they believe the retracement is concluding, thus benefitting when the stock goes back up. This same strategy is applied in bearish markets by trying to short the stock near the end of a retracement. The question is how do they know when the retracement is finishing? The Bollinger bands can help indicate this. As a reminder, a numerical way to measure where the current stock price is in the Bollinger band lines is the %B value. Again, to repeat the point of this research, investors never know for sure and hence they do not predict. They look at what the probabilities show and invest based off what the probabilities indicate.

This is why Bollinger bands are being studied in this research. SVM’s are the tools used in this study to find the effect Bollinger bands have on the probability the option expires in or out-of-the-money. The two categories of stock options in this
research are either the option expiring in-the-money or out-of-the-money. These two categories make up the two classes of the data. The fact that an option can be classified into only one of two categories is the main reason why support vector machines are being used in this research.

A certain trading technique is analyzed in this study. Bollinger bands are a measure of how over or undervalued a stock is. This is used when applying a technique using Bollinger bands. When using this technique in bullish markets, investors try to buy stocks when the current price of the stock is by the lower band. As a reminder, the lower band is two standard deviations below the simple moving average over the last twenty days. When the price of a stock reaches the lower band (which is a %B value of 0%) this tends to be a signal that the stock is becoming undervalued. Due to this fact, a technique some investors use in a bullish market is to try to use the lower band as an indicator of when a retracement is finishing and when the stock will begin to go up again [6]. This does not always happen; however, the probability the stock will go back up tends to increase as the price reaches the lower band and becomes undervalued. In bearish markets stocks can become overvalued after retracement and investors try to short stocks when they believe the stock will continue to trend down.

Another reason why this technique is argued to be effective is because using this strategy correctly can allow an investor to buy options at a discount and sell options at a higher price. The reason why this happens is because of the retracement. Consider Figure 2. The figure shows buy signals when the price reaches the lower band and sell signals when the price reaches the upper band. Because the example is showing a bullish market, the buy signals at the lower bands would be stronger signals. In bullish markets,
depending on the trend, the current price could possibly stay at or above the upper band for a period of time due to the upward trend. As a reminder of %B values, this would also mean the %B value would be greater than or equal to 100%.

Referring back to the retracement, the pullback of the stock would cause the stock to temporarily fall, which would decrease the prices of the call options and increase the prices of the put options. However, since the retracement is temporary, the lower band is a good indicator of when the pullback would conclude as shown in the figure. However, due to the fact the price temporarily begins to decrease, the call options decrease in value and the put options increase in value only temporarily. As a reminder, in a bullish market, investors make money when they buy call options and sell put options. By buying calls

Figure 2
and selling puts when the price reaches the lower band in a bullish market not only helps the probability to favor the investor, but also allows the investor to buy the call at a lower price and sell the put at a higher price. Once the price would begin to go back up, the call would then begin to become more valuable after being bought and the put would begin to decrease in value after being sold. This same strategy is also investigated for bearish markets. In a bearish market, a retracement would be a temporary increase in the stock’s price as the overall trend is down. In bearish markets, calls decrease in value and puts increase in value. The retracement would temporarily bring the call prices up and the put prices down. As the price approaches the upper band after having a downward trend, the same benefit would happen for the investor. The call options would be sold just before declining and the put options would be bought just before gaining in value. Sometimes, the market changes unexpectedly from a bullish market to a bearish market or vice versa. However, the point of this research is to prove the probabilities indicate this may happen sometimes, but overall the chances would favor the investor who applies this technique and would help him or her to sell at a higher price and buy at a lower price.

This technique is being analyzed in this study and it is being applied to options rather than to stocks themselves. Assuming this technique is a good technique and assuming that the stock price approaching the lower band is a good buy indicator of the stock in a bullish market, this would also be a good indicator to buy calls and sell puts. The reason why is because as the stock value goes up, the call options for that stock would move further in-the-money and the put options would move further out-of-the-money. In contrast, in a bearish market, if this technique proves to be effective, then when the stock becomes overvalued, this is a good indication to sell calls and buy puts.
since in this case the calls would move further out-of-the-money and puts would move further in-the-money. With this in mind, the frequency that calls and puts expire in-the-money are compared to the frequency calls and puts expire out-of-the-money in both bullish and bearish markets. The number of times the expected outcome happens in both the bullish or bearish market can determine if this strategy is a good strategy to use or whether support vector machines can be used to calculate the probability particular options expire in or out-of-the-money.
2.1: Methodology: Using Support Vector Machines

Once an investor understands the importance of probability, the question becomes how does an investor define or analyze the probability? As mentioned before, the Black-Scholes model is the most common model used to evaluate the price of options on the market. The value is based off the probability the option will expire in or out-of-the-money. This probability is based off certain factors the Black-Scholes model uses to analyze. To observe value in an option the Black-Scholes model does not observe, a certain factor is analyzed in this study that is not used in the Black-Scholes model. For this particular study, support vector machines (SVM’s) are used to analyze whether Bollinger bands have a significant effect on the probability an option will expire in or out-of-the-money. The %B value is the numerical value to measure the current price related to the Bollinger bands.

Support vector machines are learning models that analyze patterns using learning algorithms. The method first analyzes training data with data objects each of which belongs to one of two groups. The learning algorithms use the training data to develop a model, which is used to classify future data. The data and the attributes of the data are put into points on a graph. The goal of SVM’s is to construct a hyper plane, which separates the two groups of data [8]. SVM’s do not intend to minimize errors but try to build a reliable model [9].

The two most basic cases SVM’s attempt to classify are linearly separable and non-linearly separable. In a linearly separable case, a line can separate the data where all
the data in one group is on one side of the line and all the data in the other group is on the other side of the line. A non-linearly separable case is where this is not possible and accommodations must be made where certain data is incorrectly classified. When data is not linearly separable, either a kernel function is used or the soft margin approach is used which chooses a straight hyper plane that splits the data as best as possible. In the latter case, slack variables are used which estimates the error on the decision boundary. The slack variables help obtain the knowledge of how far the margin of error is in the classification [10].

SVM’s were chosen for this particular study because stock options eventually expire being in one of two groups: in-the-money or out-of-the-money. In a normal study of classification, support vector machines analyze at least two variables for each data point, plots the point on a multi-dimensional plane, observes which of the groups each object is classified in, and develops a model to classify future data. In this study, when analyzing the data, the Bollinger bands are analyzed using %B values to measure them. They are analyzed along with the expiration time of the option.

A hyper-plane can be used to separate the two groups in a linearly separable case and is sometimes used when the data is nonlinear. At other times when the data is nonlinear, more complex kernel functions might need to be used to define the separation between each group. In this particular study, the latter case is used. Both the linear and kernel approach will be discussed in the next two sections.

SVM has been proven to be highly effective in analyzing probabilities in finances [8]. SVM has often been used with stocks themselves to determine the probability a
particular trade will profit or not. In this study, SVM is used to help analyze the probabilities stock options will expire in-the-money or out-of-the-money.

2.2: Linear SVM

One of the simplest ways to perform SVM is to establish a linear function, which represents the hyper plane that separates one class from the other. Though this is one of the simplest solutions, sometimes other solutions are more appropriate. However, this type of SVM can still be used despite the data not being linearly separable, depending on the data and situation. Developing a kernel function is mostly the alternative to establishing a linear function. Kernel functions will be discussed in the next section.

For purposes of clarification, observe Figure 3 and Figure 4. Whether the data is linearly separable or not, the hyper plane is chosen based off two other hyper-planes which each pass through a particular data object or objects in their own respective classes. These points are chosen to be the points which the hyper-planes dividing the classes pass through. There is one vector chosen for each class, which are both parallel to one another. The main boundary hyper-plane is placed directly between the two vectors and is parallel to both of them. The border points on the two border vectors are the most important points in the data because these are the points which determine the support vectors. In a case where the data is linearly separable, it is easy to choose the particular data objects for the boundaries. They tend to be the points which are located the furthest away from the main group of points and closer to the points in the opposing class. In a non-linearly separable case, it is more difficult since the objective is to find which data points are correctly classified and which ones are not. The two parallel hyper-planes are
Figure 3

Figure 4

[11]

[12]
chosen based off the data objects that are considered to be correctly classified but are also considered to be located on the boundary. All the points in the training data, which are located on the opposite side of the border points, are considered to be errors. In non-linear cases, these points also determine support vectors as well as the points located on the two border hyper-planes [13].

Figure 3 is a case that is linearly separable and Figure 4 is a case that is not separable. In Figure 3, the black circles are considered one class and the white circles are considered the other. In Figure 3, notice the data objects are placed on the graph depending on the attributes $X_1$ and $X_2$. Whatever these attributes are is irrelevant for the purpose of this explanation. Notice the two border vectors each passing through the data object or data objects that are closest to the objects in the opposite class. For the class with the black circles in this case, there is only one data object but for the class with the white circles, there are two objects the border vector passes through. These three points are located on the decision boundary and are the points used to determine the hyper-plane which divides the two classes. The area between the two hyper-plane vectors is called the decision boundary [10].

This study will only apply to two dimensions to make the study simpler. For SVM’s which use two-dimensional vectors, they start with a training set of instance-label pairs $(x_i, y_i), i = 1, 2, 3, ..., m$ where $x_i \in \mathbb{R}^2$ and $y_i \in \{-1, +1\}$. This means $x_i$ is a two dimensional vector of real numbers and $y_i$ represents the group (represented by -1 or +1) the object is classified in. For linear cases of SVM, two equations are set to classify the points represented by the $x$ coordinates as shown below:
\[ w \cdot x_i - b = +1 \quad \text{for} \quad y_i = +1 \quad (2) \]
\[ w \cdot x_i - b = -1 \quad \text{for} \quad y_i = -1 \quad (3) \]

These are the two borderlines as shown in Figure 3. These are the two hyperplanes, which define the decision boundary. In a linearly-separable case, no data objects should be located in this area since this is where there is a division between the classes. One class is considered class 1 and the other -1. If a point is substituted into the equation and \( w \cdot x_i - b \geq +1 \), then the point is considered to be in class 1. If \( w \cdot x_i - b \leq -1 \), then the point is classified into class -1 [10]. This can be combined into the following equation:

\[ y_i (w \cdot x_i + b) \geq +1 \quad \forall i \quad (4) \]

The value \( w \) is also a 2 dimensional vector of real numbers, which represents the normal vector of the hyper-plane that divides the two classification groups and \( b \) is the constant, bias value. This is shown in Figure 3.

The distance between equations (1) and (2) is equal to \( 2/||w|| \) as shown in Figure 3. This equals \( \frac{2}{\sqrt{w^T w}} \). In order to maximize this, the value \( ||w|| \) or \( \sqrt{w^T w} \) must be minimized. To help make the math easier we can square it and multiply it by one half to achieve \( (1/2)w^T w \) [14]. The optimization problem is what is shown below:

\[ \arg\min_{w,b} 1/2||w||^2 \quad \text{subject to} \quad y_i (w \cdot x_i + b) \geq +1 \quad \forall i \quad (5) \]

As a reminder, the purpose of this is to find the optimal separating hyper-plane with the maximum margins on both sides. The larger the margins, the more definitive each class is. Due to the need to find the stationary point this becomes what is called a
Lagrange multiplier, which is a way of finding the maxima and minima of a function using partial derivatives. The problem becomes what is shown below:

$$L_p(w, b, \alpha) = \left(\frac{1}{2}\right)w^T \cdot w - \sum_{i=1}^{n} \alpha_i(y_i(w \cdot x_i) + b) - 1$$  \hspace{1cm} (6)

This means we will be taking the partial derivative of this equation with respect to the vector $w$ and the value $b$. The purpose is to find the minimum value of $L_p$ with respect to $w$ and $b$. The equation must also be maximized with respect to $\alpha$, which is the Lagrange multiplier. To find optimal values, the partial derivative with respect to each critical value is found and then the resulting equation of each derivative is set equal to 0 [10]. Once this is done, the equation is solved to find the value of the variable for the optimal condition. Each partial derivative is done for each variable. The first derivative will be the derivative with respect to $w$. The resulting equation is then set equal to zero as shown below:

$$0 = w - \sum_{i=1}^{n} \alpha_i(y_i(x_i))$$  \hspace{1cm} (7)

Therefore, by doing this, we are able to solve the equation for the vector $w$ to find what $w$ equals to receive the best result. Simply solve for $w$ by performing simple algebra.

$$w = \sum_{i=1}^{n} \alpha_i(y_i(x_i))$$  \hspace{1cm} (8)

The second partial derivative is with respect to $b$. It is also set equal to zero:

$$\sum_{i=1}^{n} \alpha_i(y_i) = 0$$  \hspace{1cm} (9)

With this done, it is now necessary to maximize the function by substituting the optimal value for $w$ into the equation. To be clear, the $w$ in equation (6) is now substituted with what $w$ equals in equation (7), which results in the following equation:
\[ L = \left( \frac{1}{2} \right) \sum_{i=1}^{n} a_i(y_i(x_i)) \cdot \sum_{j=1}^{n} a_j(y_j(x_j)) - \sum_{i,j=1}^{n} a_iy_ix_ja_jy_jx_j - \sum_{i=1}^{n} a_iy_i + \sum_{i=1}^{n} a_i \]  

(10)

By using equation (9) to set \( \sum_{i=1}^{n} a_i(y_i) \) equal to 0 and by simplifying the equation, the optimal equation becomes simplified to the equation shown below:

\[ L = \sum_{i=1}^{n} a_i - \left( \frac{1}{2} \right) \sum_{i,j=1}^{n} a_iy_ix_ja_jy_jx_j \]  

(11)

To handle the inequality constraints, the Lagrange multipliers have two conditions known as the Karush-Kuhn-Tucker conditions:

\[ a_i \geq 0 \]  

(12)

\[ a_i(y_i(w \cdot x_i + b) - 1) = 0 \]  

(13)

Using all these equations and conditions, we can find the separable line. The Lagrange multipliers would have to be found using quadratic programming. Once these are found, we can use these equations to develop the best hyper-plane using the support vectors of the particular data set [10]. This will be discussed in a later section.

Unfortunately, it is not always the case where the data is linearly separable. This means in particular data sets, there is no way to draw a straight line which separates all the data in one class from all the data in the other class. Again, this is referred to as a case that is not linearly separable. However, support vector machines can still be used in this case where it can define a hyper plane. If the data is not linearly separable, then using a linear strategy will always result with a function that has errors. This means some of the data objects, which are contained in one class, are incorrectly classified by the hyper-plane into the other class. The goal is to try to create a reliable model. This is done by
constructing a hyper plane in the data which has the fewest number of data objects that are incorrectly classified [10].

In this particular study, no cases are linearly separable due to a variety of factors such as arbitrage. But SVM can help observe whether certain values of %B will very likely cause an option to expire in-the-money and values which will very likely cause an option to expire out-of-the-money. As a reminder, nothing is guaranteed with the market or with options. Instead, probabilities are used to help an investor make money in the long run. Due to nothing being guaranteed, there are cases where an option contains particular attributes which would theoretically cause an option to expire in-the-money, however, arbitrage causes it to actually expire out-of-the-money or vice versa. Arbitrage causes the market to do temporary, unexpected fluctuations; however, applying a strategy based off probabilities can help a trader minimize risk and make money overall despite losing money in some trades [15].

The case in this study is such a case, which is shown in Figure 4, but the %B value is tested to see if it has a significant effect on the probability of the outcome even when it will not guarantee a particular outcome. As a reminder, this does not mean the %B value can ultimately determine the outcome of a particular stock option, but rather the question is if it can reveal a particular probability the Black-Scholes Model does not see or take into consideration. This is a way of finding value (or lack of value) in an option that others cannot see.

In a non-linearly separable case, another variable is introduced in these equations: $\xi$. This variable is a slack variable to represent the tolerance of misclassification. In a non-separable case, Equations (2), (3), and (4) become the following: [16]
\[ w \cdot x_i - b = +1 - \xi \text{ for } y_i = +1 \] (14)

\[ w \cdot x_i - b = -1 - \xi \text{ for } y_i = -1 \] (15)

\[ y_i(w \cdot x_i + b) \geq +1 - \xi \; \forall i \] (16)

The slack variable estimates the error of the decision boundary on the particular data object it is associated with. If a line parallel to the decision boundary runs through the data object associated with each slack variable, the distance between that line and the decision boundary would be \( \xi / ||w|| \). For example, observing Figure 4, there is a slack variable \( \xi_1 \), associated with a data object that would be incorrectly classified as a red star and is instead a blue circle. If a line parallel to the decision boundary passed through that particular data object, the distance between that line and the boundary is \( \xi_1 / ||w|| \). We want to try to minimize the number of slack variables and the quantity of the slack variables as much as possible but still have a definitive boundary, which separates the data objects in one class from the data objects in the other. Sometimes this might require narrowing the margin of the boundaries. To accommodate for the error and to try to minimize the slack variables, the following equation is given modified from Equation (5):

\[ f(w) = \frac{||w||^2}{2} + C (\sum_{i=1}^{N} \xi_i)^k \] (17)

This also causes a change in Equation (6) to the following:

\[ L_p(w, b, \xi, \alpha, \mu) = \frac{||w||^2}{2} + C (\sum_{i=1}^{N} \xi_i)^k - \sum_{i=1}^{n} \alpha_i(y_i(w \cdot x_i + b) - 1 + \xi_i) - \sum_{i=1}^{N} \mu_i \xi_i \] (18)

When taking the partial derivatives with respect to \( w \) and \( b \), Equation (7) and Equation (8) remain the same with the partial derivatives being taken in respect to the
same variables. By taking the partial derivative with respect to $\xi_i$, another equation is provided for this case to find what the optimal value of the variable $C$ is:

$$ 0 = C - \alpha_i - \mu_i $$

(19)

With simple algebra, the variable $C$ is given:

$$ C = \alpha_i + \mu_i $$

(20)

By substituting all these optimization values for the variables into Equation (18), we arrive at the following equation:

$$ L_p(w, b, \alpha) = \left(\frac{1}{2}\right) \sum_{i,j=1}^{n} \alpha_i y_i (\alpha_j y_j x_j) + C \sum_{i=1}^{n} \xi_i - \sum_{i=1}^{n} \alpha_i (y_i \sum_{i=1}^{N}(\alpha_j y_j x_j + b) - 1 + \xi_i) - \sum_{i=1}^{N}(C - \alpha_i)\xi_i $$

(21)

Though this is a new optimization model with new variables, the solved equation becomes the same equation as Equation (11):

$$ L = \sum_{i=1}^{n} \alpha_i - \left(\frac{1}{2}\right) \sum_{i,j=1}^{n} \alpha_i y_i (\alpha_j y_j x_j) $$

(22)

However, there is a subtle, yet important difference between Equation (11) in the linearly separable case and Equation (22) in the non-linearly separable case. There are more necessary equality constraints that are transformed from the inequality constraints [10]. The constraints are shown below:

$$ \xi_i \geq 0, \quad \alpha_i \geq 0, \quad \mu_i \geq 0 $$

(23)

$$ \alpha_i (y_i (w \cdot x_i + b) - 1) = 0 $$

(24)

$$ \mu_i \xi_i = 0 $$

(25)

In the separable case, a constraint on Equation (11) requires $\alpha_i$ to be greater than or equal to 0. In the non separable case Equation (22) gives a constraint on $\alpha_i$ that
requires the variable to not only be greater than or equal to 0 but also to be less than or equal to the variable $C$ [14].

As is the linearly separable case, using all these equations, we can find the separable line and the Lagrange multipliers would have to be found using quadratic programming. Notice an extra Lagrange multiplier ($\mu_i$) is added in the non-separable case. In this case, some data objects are misclassified and there is some tolerance for errors; however, in certain situations, this type of method can be proven effective once the Lagrange multipliers are found using quadratic programming. Before discussing quadratic programming, it is important to mention the alternative to developing a linear hyper-plane for the support vectors.

2.3: Kernel Methods

Kernel methods are another effective way to develop a function to separate the data. In certain situations, a straight line cannot accurately divide the classes of data. In some cases, data is not at all separable and must take errors into account. However, even when data is not at all linearly separable, it might be separable in other ways.

Observe Figure 5 and Figure 6 to illustrate a point. Figure 5 shows a case where a line that is not straight is the best line to define the data. In this case, it might be legitimate to make it simple and define a straight line, which has some errors, however, the function of the line in Figure 5 is more accurate. For Figure 6, it is definitely not a case to try and separate the data with a straight line. Figure 6 is not at all linearly separable but it is definitely separable in another way. Since one class is grouped in a
Figure 5

Figure 6
circle and the other class is grouped outside of it, the kernel function in this case would be a radial basis function.

As shown in both Figure 5 and Figure 6, when kernel functions are developed, they can help develop a reliable model based off the training data. The model is used to develop the boundary. This boundary is defined using linear SVM after the kernel function transforms the space. A kernel function is mostly represented by the variable, \( \Phi \) [14].

The purpose of the kernel function is to map the data from the original coordinate space to a new space. The data is mapped in such a way so a linear decision boundary can be used to separate the classes [10]. Normally, when mapping the data from one space to another, \( XI \) is defined as the data in the original space and \( XI' \) is the space \( XI \) is mapped to. Once the data is mapped into a new space, the kernel function takes the inner product of the two new vectors. Sometimes, to transform the space into a space that has linearly separable classes, the transformed space can have more dimensions than the original space. For example, the following transformation is a legitimate transformation from one space to another using the kernel method:

\[
\Phi : (X_1, X_2) \rightarrow (X_1^2, X_2^2, \sqrt{2}X_1, \sqrt{2}X_2, 1)
\]

This would cause the vector \( w \) to be a five dimensional vector instead of two, therefore, the equation resulting from the dot product would be the following: [10]

\[
W_4X_1^2 + W_3X_2^2 + W_2\sqrt{2}X_1 + W_4\sqrt{2}X_2 + W_0 = 0
\]

A kernel function encompasses two functions and multiplies the result of both functions. The symbol \( \Phi(x_i) \) represents the result of the mapping of \( x_i \) from one input.
space to another. The following represents the symbol of the kernel function by mapping
the spaces of \( x1 \) and \( x2 \) and performing the dot product of the two results:

\[
K(x1, x2) = \Phi(x1) \cdot \Phi(x2)
\]  

(26)

The computation of a kernel function obtains a value which represents the
similarity between the two data points. In cases where the original space transforms into a
space that has many dimensions, this can become useful. It is a way of expressing the dot
product of the transformed space in terms of the original space. This is referred to as the
kernel trick and can be used to efficiently express the data when the transformed space
has many dimensions [10]. The three most common kernel functions are the three below:

Polynomial Kernel: \( K(x1, x2) = (1 + x1 \cdot x2)^d \)  
Radial Basis Function Kernel: \( K(x1, x2) = e^{(x1- x2)^2 / (2\sigma^2)} \)  
Sigmoid Kernel: \( K(x1, x2) = \tanh (kx1 \cdot x2 - \delta) \)  

With the kernel function, the equations from the linear case remain very similar,
except where there was normally the \( x \) vector. With the kernel method, the kernel
functions are put in place of the \( x \) vector in the linear case. Referencing Equation (2) and
Equation (3) the equations with kernel functions would change to the following:

\[
w \cdot \Phi(x_i) - b \geq +1 \quad \text{if} \quad y_k = 1
\]  

(30)

\[
w \cdot \Phi(x_i) - b \leq -1 \quad \text{if} \quad y_k = -1
\]  

(31)

As in the linear case, this would be equivalent to the following:

\[
y_i (w \cdot \Phi(x_i) + b) \geq +1 \quad \forall i
\]  

(32)
As in the linear case, if the result of the function is greater than or equal to 1, the data object is classified in class 1, if it results in less than or equal to −1, it is classified into class −1 [16]. The other equations are also the same with \( \Phi(x_i) \) replacing \( x_i \): [10],

\[
w = \sum_{i=1}^{n} \alpha_i (y_i (\Phi(x_i)))
\]

(33)

\[
L_p(w, b, \alpha) = \left( \frac{1}{2} \right) w^T \cdot w - \sum_{i=1}^{n} \alpha_i (y_i (w \cdot \Phi(x_i) + b) - 1)
\]

(34)

\[
L = \sum_{i=1}^{n} \alpha_i - \left( \frac{1}{2} \right) \sum_{i,j=1}^{n} \alpha_i \alpha_j y_i y_j K(x_i, x_j)
\]

(35)

Equation (34) and (35) will have the constraint represented by Equation (12) as well as the following:

\[
\alpha_i (y_i (w \cdot \Phi(x_i) + b) - 1) = 0
\]

(36)

Sometimes after transforming the space, the data still does not become linearly separable. The equations dealing with the linearly non-separable case are applied when the kernel function transforms the space and the data is not linearly-separable in the resulting space. Again, the same equations are used with the slack variable \( \xi \) estimating the error of the decision boundary as before and changing \( x_i \) to \( \Phi(x_i) \): [16]

\[
w \cdot \Phi(x_i) - b = +1 - \xi \quad \text{for} \quad y_i = +1
\]

(37)

\[
w \cdot \Phi(x_i) - b = -1 - \xi \quad \text{for} \quad y_i = -1
\]

(38)

\[
y_i (w \cdot \Phi(x_i) + b) \geq +1 - \xi \quad \forall i
\]

(39)

\[
L_p(w, b, \xi, \alpha, \mu) \|w\|^2 + C (\sum_{i=1}^{N} \xi_i) + \sum_{i=1}^{n} \alpha_i (y_i (w \cdot \Phi(x_i) + b) - 1 + \xi_i) - \sum_{i=1}^{N} \mu_i \xi_i
\]

(40)

\[
L = \sum_{i=1}^{n} \alpha_i - \left( \frac{1}{2} \right) \sum_{i,j=1}^{n} \alpha_i \alpha_j y_i y_j K(x_i, x_j)
\]

(41)
For Equation (39) and (40), the constraints are represented by Equations (23) and (25) as well as the following:

$$\alpha_i(y_i(w \cdot \Phi(x_i) + b) - 1) = 0$$  \hspace{1cm} (42)

Equation (36) and Equation (42) are similar to Equations (13) and (24) respectively. They simply change the data object from being in the original space ($x_i$) to being in the transformed space ($\Phi(x_i)$).

As is the case for performing linear SVM, quadratic programming solves for the Lagrange multiplier $\alpha_i$ (and for the Lagrange multiplier, $\mu_i$ in the non-separable case), which allows the parameters $w$ and $b$ to be solved when the solved value is put into the equation. The following section addresses how to do this.

2.4: Quadratic Programming

Quadratic programming involves finding the optimization values given certain constraints. The process involves both linear algebra and differential calculus. There are different operations of quadratic programming but all methods are used using computer programs. For this study a computer program is written in Java using pre-defined classes and methods which performs the mathematical operations.

In the process of quadratic programming, all the Lagrange multipliers ($\alpha_i$ and $\mu_i \forall i$) associated with the support vectors are found. There is one Lagrange multiplier associated with each data object in the data, however, any Lagrange multiplier associated with any data object that is not a support vector, is equal to 0. Quadratic programming is used to find the Lagrange multipliers associated with the support vectors. Once all the support vector Lagrange multipliers are found, their values can be substituted in Equation
(8) in the linear case and in Equation (33) in the kernel case for each support vector. This allows the vector $w$ to be calculated. Once the value of $w$ is found, the value can be substituted in Equation (13) when the data is linearly separable and the linear SVM is used, in Equation (24) when the data is not linearly separable but the linear SVM is used, or in Equation (42) when using the kernel method. The vector $w$ and the Lagrange multipliers are substituted in one of those three equations. Using algebra, the value of $b$ can be solved for each dimension. Once the value of $b$ is found for each dimension of the vectors, the value of $b$ is averaged between each dimension. Once this is found, the values of $w$ and $b$ can be substituted in the equations to define the boundary and the margin between the classes of data. The training data gives the necessary information to develop these equations. The equations are then used for future data to try and correctly classify it as much as possible.

For this particular study, the Bollinger bands are one dimension with the time until expiration being the other dimension.

2.5: Previous Uses of Support Vector Machines in Finances

Support vector machines have been a great tool to use analyzing finances [8]. Their categorization of data and the algorithm it uses for predicting data has proven to be effective.

A previous study by Manish Kumar and M. Thenmozhi proved support vector machines to be an effective tool to predict financial time series and forecasting stock index movement. It was proven to outperform random forest, neural network, and other models [8].
SVM’s have also been used to analyze credit applicants. A study done by KB Schebesch and R Stecking proved SVM’s to be successful at detecting good credit versus bad credit. In the study, it was proven that the SVM’s were able to establish critical, well defining decision rules to forecast future data based off the training data [13]. Jan-Henning Trustorff, Paul Markus Konrad, and Jens Leker also did a similar study by trying to predict credit risk. The study proved that using SVM’s can classify very accurately due to the learning algorithms that the method consists of [16].

SVM’s have also been successfully used to analyze and correctly classify financial assets. Pankaj Gupta, Mukesh Kumar Mehlawat, and Garima Mittal performed a study to use SVM’s to classify financial assets into three distinct classes by evaluating return, risk, and liquidity. The approach was proven to provide successful, and accurate results.

SVM’s were first developed in the 1990’s. Despite SVM’s not being known for very long, SVM’s have proven to be very useful in many studies involving classification of data. Due to it being proven to be effective and efficient at classifying data, it is used in this study to predict options expiring in or out-of-the-money.
Chapter 3
Tests And Analysis

3.1: Parameters and Description of Tests

Both bullish and bearish markets are studied and both call options and put options are also studied. There are two main cases that can cause a model to be incorrect: under-fitting and over-fitting. Under-fitting is where a model does not have enough data to draw an accurate model and conclusion. Over-fitting is the opposite where the model represents the sample too much and though might have a very detailed model and accurate model for the particular sample, it does not have an accurate model for the overall population from which the sample is taken from. The two kernels used in this study are the polynomial kernel and the radial basis kernel. The polynomial kernel is used the most often to prevent over-fitting with more complicated kernels and yet prevent under-fitting with very simple, general kernels such as the linear model. The radial basis kernel is still used but not as often since it seems to cause over-fitting.

The parameters are shown in Figure 7. The kernel type is the type of kernel and the degree only applies to the polynomial kernel, which represents the $d$ variable for Equation (27). The parameter, $svm\_type$ is the type of support vector machines used which is always C-SVM which is the most common type used and is what has been described in the previous sections. The gamma variable is what $1/2\sigma^2$ equals in the radial basis function. The $nu$ variable is irrelevant since it only applies to other support vector machines. The $cache\_size$ is the size of the memory of the cache. The $eps$ is the tolerance of termination and the $p$ parameter is also irrelevant due to being used only by another support vector machine. The probability parameter being set to 1 allows the
program to obtain probability values for all particular cases. This will not be analyzed in
this study as only the model itself is analyzed.

/**
 * Constructor which sets the default values of the parameters of svm_parameter
 */
public SupportVectorMachine()
{
    parameter = new svm_parameter();

    //Provide default values of the variables
    parameter.svm_type = svm_parameter.C_SVC;
    parameter.kernel_type = svm_parameter.POLY;
    parameter.degree = 4;
    parameter.gamma = 0.175;
    parameter.nu = 0.5;
    parameter.cache_size = 20000;
    parameter.C = 1;
    parameter.eps = 0.001;
    parameter.p = 0.1;
    parameter.probability = 1;

    problem = new svm_problem();
}

Figure 7

3.2: Test Results

Figure 8 shows data for S&P500 call options. These are options from March 2013
to March 2014. The graph shows data for a stock on a particular day. It shows what
would have resulted if a particular call or put option was bought or sold on a particular
day. Blue data points represent options that eventually expired out-of-the- money and red
data points represent options that expired in-the-money. The Y axis shows the percent B values for the stock and the X axis shows the days until expiration for each option. In other words, if a point on the graph is located at (30, 50), then that particular option was able to be either bought or sold 30 days from expiration and the percent B value for the stock that option represents was 50% on that day. All data in a particular study have one thing in common: they all have the same $\delta$. All data in a particular study have a value for $\delta$ of either 0.2 or 0.3. As a reminder, the value, $\delta$, for all options, gives the change in price of an option in relation to every dollar increase in value of the corresponding stock. Theoretically, the value $|\delta|$ for a particular option at a particular strike price also represents the probability that option will expire in-the-money. Stock options which have a value for $|\delta|$ of less than 0.5 ($|\delta|$ due to put options having a negative value for $\delta$) are currently located out-of-the-money and options which have a value for $|\delta|$ of greater than 0.5 are currently located in-the-money. Therefore, all options used in these studies are located out-of-the-money at the time they are analyzed. The data points that are blue remained out-of-the-money at expiration and data points that are red expired in-the-money on the day of expiration. The lower the $|\delta|$, the further out-of-the-money the particular option is and the higher the value for $|\delta|$ the further in-the-money it is. The value cannot be less than 0 but cannot exceed 1. With the values staying the same, this means the strike prices for all these options differ. To understand this, observe both Figure 9 and Figure 10. Figure 9 shows the option chain for both call and put options on the S&P500 on March 28 2013. These options expire on April 19, 2013. The call options are located on the left and the put options are located on the right. Options that are highlighted are in-the-money and the others are out-of-the-money. Notice that if a call
and put option both have the same strike price, one must be out-of-the-money while the other must be in-the-money. Notice the options that have a value for $\delta$ equal to or close to 0.2 and -0.2 respectively are highlighted. The call option with a value for $\delta = 0.2$ has a strike price of 1595 and the put option with $\delta = -0.2$ value has a strike price of 1525. Figure 10 shows the option chain for April 1, 2013. Notice that because the price of the
S&P500 declined, the value $\delta$ for the call option at a strike price of 1595 (as well as the other call options) and the value for the put option at a strike price of 1525 (as well as the other put options) decreased. For April 1st, the strike prices that have a value for $\delta$ of 0.2

![Figure 9](image-url)

and -0.2 are now 1590 and 1520. In this study, because the options observed have the same $\delta$ value but are on different days, the strike prices differ. The reason why the values for $\delta$ for all these options are the same for a particular study is because $\delta$ gives an
indication of how far in or out-of-the-money a stock option is. It takes certain effects in the market into account including but not limited to the life of the option, the volatility of the market, and the type of market (bullish or bearish). Call options increase in value in bullish markets and put options increase in value in bearish markets. With all this in mind, Figure 8 shows all the S&P 500 calls at a value for $\delta$ equal to 0.2 from March 2013 to March 2014. Data points higher on the $Y$ axis show options whose stock had higher $\%B$ values on the day they were bought or sold and data points further to the right have

<table>
<thead>
<tr>
<th>Last</th>
<th>Net.Chng</th>
<th>Volume</th>
<th>Open</th>
<th>High</th>
<th>Low</th>
</tr>
</thead>
<tbody>
<tr>
<td>1562.17</td>
<td>-7.02</td>
<td>0</td>
<td>1566.18</td>
<td>1570.37</td>
<td>1538.47</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Calls</th>
<th>Delta</th>
<th>Gamma</th>
<th>Theta</th>
<th>Bid</th>
<th>Ask</th>
<th>Vega</th>
</tr>
</thead>
<tbody>
<tr>
<td>APR 13</td>
<td>.90</td>
<td>.00</td>
<td>-.16</td>
<td>.39</td>
<td>51.00</td>
<td>55.90</td>
</tr>
<tr>
<td>APR 13</td>
<td>.83</td>
<td>.01</td>
<td>-.30</td>
<td>.86</td>
<td>46.60</td>
<td>54.90</td>
</tr>
<tr>
<td>APR 13</td>
<td>.80</td>
<td>.01</td>
<td>-.33</td>
<td>.96</td>
<td>41.50</td>
<td>50.00</td>
</tr>
<tr>
<td>APR 13</td>
<td>.78</td>
<td>.01</td>
<td>-.34</td>
<td>1.01</td>
<td>37.50</td>
<td>46.80</td>
</tr>
<tr>
<td>APR 13</td>
<td>.76</td>
<td>.01</td>
<td>-.33</td>
<td>1.06</td>
<td>31.20</td>
<td>41.50</td>
</tr>
<tr>
<td>APR 13</td>
<td>.75</td>
<td>.01</td>
<td>-.33</td>
<td>1.11</td>
<td>28.90</td>
<td>34.90</td>
</tr>
<tr>
<td>APR 13</td>
<td>.79</td>
<td>.01</td>
<td>-.39</td>
<td>1.23</td>
<td>26.00</td>
<td>34.60</td>
</tr>
<tr>
<td>APR 13</td>
<td>.85</td>
<td>.01</td>
<td>-.39</td>
<td>1.28</td>
<td>25.90</td>
<td>27.00</td>
</tr>
<tr>
<td>APR 13</td>
<td>.91</td>
<td>.01</td>
<td>-.39</td>
<td>1.33</td>
<td>21.50</td>
<td>22.20</td>
</tr>
<tr>
<td>APR 13</td>
<td>.96</td>
<td>.01</td>
<td>-.40</td>
<td>1.36</td>
<td>18.60</td>
<td>19.90</td>
</tr>
<tr>
<td>APR 13</td>
<td>.51</td>
<td>.01</td>
<td>-.39</td>
<td>1.38</td>
<td>15.60</td>
<td>18.80</td>
</tr>
<tr>
<td>APR 13</td>
<td>.46</td>
<td>.01</td>
<td>-.38</td>
<td>1.37</td>
<td>12.60</td>
<td>13.80</td>
</tr>
<tr>
<td>APR 13</td>
<td>.40</td>
<td>.01</td>
<td>-.35</td>
<td>1.34</td>
<td>11.40</td>
<td>12.20</td>
</tr>
<tr>
<td>APR 13</td>
<td>.35</td>
<td>.01</td>
<td>-.33</td>
<td>1.28</td>
<td>7.60</td>
<td>8.80</td>
</tr>
<tr>
<td>APR 13</td>
<td>.29</td>
<td>.01</td>
<td>-.30</td>
<td>1.19</td>
<td>5.70</td>
<td>6.80</td>
</tr>
<tr>
<td>APR 13</td>
<td>.24</td>
<td>.01</td>
<td>-.26</td>
<td>1.07</td>
<td>4.10</td>
<td>4.90</td>
</tr>
<tr>
<td>APR 13</td>
<td>.19</td>
<td>.01</td>
<td>-.23</td>
<td>.95</td>
<td>2.95</td>
<td>3.90</td>
</tr>
<tr>
<td>APR 13</td>
<td>.14</td>
<td>.01</td>
<td>-.19</td>
<td>.79</td>
<td>2.05</td>
<td>2.80</td>
</tr>
<tr>
<td>APR 13</td>
<td>.11</td>
<td>.01</td>
<td>-.16</td>
<td>.67</td>
<td>1.60</td>
<td>1.90</td>
</tr>
<tr>
<td>APR 13</td>
<td>.08</td>
<td>.00</td>
<td>-.12</td>
<td>.51</td>
<td>.75</td>
<td>1.35</td>
</tr>
<tr>
<td>APR 13</td>
<td>.06</td>
<td>.00</td>
<td>-.10</td>
<td>.41</td>
<td>.50</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Figure 10
more days until expiration. In this particular case all the options have a life of 56 days or less. This was a bullish market for the S&P500 during the time. As shown in the data, it
is very easy to see that more options expire out-of-the-money that have higher \%B values and more options expire in-the-money that have lower percent B values despite the fact they are bought or sold out-of-the-money at a value for $\delta$ equal to 0.2.

With all this data, a model can be developed to determine how it predicts future data. Figure 11, Figure 12, and Figure 13 all use a polynomial kernel. Figure 11 has a
degree of 2, Figure 12 degree 3, and Figure 13 degree 4. Figure 14 shows a model developed using the RBF kernel.

The data for the first has been taken manually. Because of this, this was causing the test to be very time consuming. Therefore, after developing the model, some
improvements have been made to allow the data to be analyzed automatically and more efficiently. The three polynomial models (Figure 12, 13, and 14) each very clearly indicate the effects of Bollinger bands as was originally hypothesized. All three polynomial graphs are showing that when the stock for the corresponding option is lower on the Bollinger bands, there is a higher chance the option will expire in-the-money and if it is higher, there is a higher chance the option will expire out-of-the-money. The radial
basis kernel in Figure 14 that is developed from the data also shows this but not as clearly as the polynomial graphs due to it being in more detail.

Models were developed from this data, however, the program has been improved on to help analyze and observe data more efficiently. As the program has been improved on, three months of data has been tested and the fourth month is used to test the accuracy.
of the model. The data for the first test for the bull market is shown in Figure 15 and shows the June, July, and August 2013 Calls for the S&P 500. Before observing the models developed from the support vector machines, observe figures 16, 17, 18, 19, 20, and 21. These graphs show very clearly why these tests are being conducted. There are three bar graphs which are shown in Figures 16, 17, and 18. Figure 16 shows how many options were in and out-of-the-money altogether. Figure 17 is showing all options on

![Figure 16](image_url)
days where the S&P 500 had a %B value greater than 50% and Figure 18 shows the options where the index for that day had lower than a 50% B value. After originally observing all the options in Figure 16, notice in Figure 17 how many of those options expired out-of-the-money and few expired in-the-money. In contrast, notice in Figure 18 how many of the options expired in-the-money and fewer expired out-of-the-money. Observations such as these are why these tests are being conducted. In these tests, it is
first observed whether the support vector machines create a model that support these bar graphs. Figure 19, 20, and 21 show this more clearly. These figures are pie graphs showing the percentage and portion of options in-the-money and out-of-the-money. Figure 19, 20, and 21 show this more clearly. These figures are pie graphs showing the percentage and portion of options in-the-money and out-of-the-money. Figure 19 shows the portion of options in and out-of-the-money altogether. Figure 20 shows the portion of options that expired in and out-of-the-money when the S&P500 index for that particular
day had greater than a 50 %B value. Figure 21 shows all the options with a %B value less than 50%. It is expected for the portion of options that expire out-of-the-money overall to be around 80% and the portion of options that expire in-the-money to be around 20% since all these options have a value for $\delta$ equal to 0.2. However, when observing the other two figures, the probabilities change significantly due to the different %B values.

Figure 19
Figure 20

Options In Or Out of the Money

Out of the Money (90.6%)  
In the Money (9.4%)

Figure 21

Options In Or Out of the Money

Out of the Money (56.3%)  
In the Money (43.8%)
In Figure 20, which shows higher %B values, observe that the probability of choosing an option with $\delta = 0.2$ expires out-of-the-money increases to over 90% and the probability the option expires in-the-money decreases to less than 10%. However, when the %B values are less than 50%, the probability the option expires out of the money decreases to 56% and the probability the option expires in-the-money increases to almost 44%. The combined options in Figure 20 and Figure 21 are the same options in Figure 19. They divide up the options according to whether the Bollinger bands are greater than or less than 50%. With this in mind, this shows the significance of Bollinger bands and how they can greatly affect the probability of the results of the options. This is why this is being analyzed in further detail with support vector machines. As previously mentioned, the support vector machines analyze the Bollinger bands and the days until expiration of each option for three months and develop a model to predict the fourth month. Comparing the predictions with the actual results of the fourth month then tests the model.

The first study is on stock options for the S&P500. The value for $\delta$ for all the calls in the first study is 0.2. All the options in this study expire in 56 days or less. Figure 22 shows an SVM model using a polynomial kernel with a degree equal to two. This first test supports the original hypothesis due to the lower %B values predicted to expire in-the-money and the higher %B values predicted to expire out-of-the-money. The only exception is the prediction of very high %B values shortly before expiration of the option. This is probably developed due to some of the few options in the data expiring in-the-money when the market was increasing very rapidly shortly before expiration. This model is developed based off the data in Figure 15.
Figure 22

Figure 23

52
Figure 24

Figure 25

53
The model developed in Figure 23 is also a polynomial kernel but with a degree of 3 instead of 2 as in Figure 22. The model in Figure 23 is also very supportive of the original hypothesis. This particular model predicts if the %B value is above 65 to 70%, most likely the option will expire out-of-the-money and if the stock has a lower %B value than that, the option will most likely expire in-the-money.

One more polynomial kernel was used with a degree of four to develop the model shown in Figure 24. This model mostly showed the affects of the number of days until expiration since it shows the probability of being in-the-money increasing moving further away from the day of expiration. Nevertheless, the model in Figure 24 does show some support for higher %B values indicating a lower chance of the option expiring in-the-money. Even though the red in the graph is to the right, the majority of the area inside the red is below the 50% %B value.

The radial basis kernel is the final kernel used for this test to develop a model. However, it is very difficult to see a trend from this model due to over-fitting. All four of these kernels are used again in another test where $\delta = 0.3$. The same options for the same days are studied. Therefore, the days until expiration and the Bollinger bands are the same. The only difference is more options expire in-the-money since the options are not located as far out-of-the-money on the particular day they are analyzed. The data is shown in Figure 26 and the tests are shown in Figures 27, 28, 29, and 30 respectively.

After analyzing the calls, the put options are also analyzed in the exact same way on the same dates, same expiration dates and prices, and the same values for $\delta$ (0.2 and 0.3). The data for the 0.2 put options from June to August 2013 is shown in Figure 31 and the tests in Figure 32, 33, 34, and 35 are the models developed using the same kernels.
Figure 26

Figure 27
Figure 28

Figure 29
Figure 30

Figure 31
Figure 32

Figure 33
Figure 36

Figure 37
Figure 38

Figure 39
The same data is analyzed with $\delta = 0.3$ and shown in Figure 36. The tests using the same kernels for this data are shown in Figures 37, 38, 39, and 40.

Certain test results are very interesting such as the one shown in Figure 32 using a polynomial kernel with a degree of 2. This test shows that in any case (no matter what the %B values are), the probability the put option expires out-of-the-money is greater than the probability of any expiring in-the-money due to the bullish market. However, as expected, the majority of the tests show the chance of options expiring in-the-money increasing with higher %B values and the chance of options expiring out-of-the-money increasing with lower %B values. There are some models that show certain extreme cases such as in Figure 37 which show very high %B values shortly before expiration where
the probability the option expires out-of-the-money is greater than it expiring in-the-money in that situation. However, overall, the models support the general hypothesis.

The results of testing all these models are shown in Table 1. The prediction of the put options are more accurate than the prediction of the call options. What might have some influence in this is the fact that most put options expire out-of-the-money in bullish markets. But something to take note is that many of the models developed that support the original hypothesis have been the most accurate such as the model developed from a polynomial kernel of degree 3 for the call options with $\delta = 0.2$. The models developed for the 0.3 calls are not as accurate but are also not as supportive of the hypothesis as the models developed for the 0.2 calls. Also, the further out-of-the-money the options are at

<table>
<thead>
<tr>
<th>Calls/Puts</th>
<th>$\delta$</th>
<th>Kernels</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>RBF</td>
</tr>
<tr>
<td></td>
<td></td>
<td>d = 2</td>
</tr>
<tr>
<td>Calls</td>
<td>0.2</td>
<td>47.4%</td>
</tr>
<tr>
<td>Calls</td>
<td>0.3</td>
<td>50%</td>
</tr>
<tr>
<td>Puts</td>
<td>-0.2</td>
<td>92.1%</td>
</tr>
<tr>
<td>Puts</td>
<td>-0.3</td>
<td>89.5%</td>
</tr>
</tbody>
</table>

Table 1
the time of analysis, the easier it is to develop a model. Though the put options are predicted much more accurately, this can largely be due to the fact that few puts expire in-the-money in a bullish market. The unbelievable accuracy for the put options can be an indicator that the particular market is bullish. Nevertheless, for both call and put options, it is proven that with certain kernels in certain situations, support vector machines can be a very useful tool to analyze the probabilities and predict the results of stock options.

After looking at the bullish market for the S&P500, a bearish stock analyzed in this study is the Peabody Energy stock (Symbol is BTU). Call options are not analyzed for the bearish market since so many call options expire out-of-the-money in a bearish market. Figure 41 shows the data for the put options with $\delta = 0.2$ from September to November. Figure 42, 43, 44, and 45 show the models for the data. Figure 46 shows the put options with a value for $\delta$ equal to 0.2 from September to November. Figure 47, 48, 49, and 50 show the models for the data with the 0.3 $\delta$. The same tests with the same kernels as is done with the S&P500 is done for the Peabody Energy stock. As a reminder, with these figures showing a bearish market with put options, the exact opposite is expected in contrast to call options. The higher %B values are expected to increase the probability the put options expire in-the-money whereas the probability decreases with lower %B values. One fact to take into consideration is the fact that bearish markets are more volatile than bullish markets and the market moves much faster in bearish markets than bullish markets.
Figure 43

Figure 44

66
Figure 45

Figure 46
Figure 47

Figure 48

68
Figure 49

Figure 50
Again, many of the models (but not all) support the original hypothesis. Other than the polynomial kernel developed for the $\delta = 0.2$, the tests are very accurate. Also the polynomial kernel with $\delta = 0.2$ created a model exactly the opposite from what was expected which helps explain its inaccuracy. The results are shown in Table 2. As in the bullish market, the support vector machines have done very well overall at predicting the option results based off previous data. The options predicted are the BTU put options that expire 56 days or less for December.

However, one word of caution on using this strategy for the bearish market, the range of accuracy is very wide. This is probably due to the fact that bearish markets are much more volatile than bullish markets. Therefore, the results below might indicate that it is better to use this strategy in bullish markets than bearish markets.

<table>
<thead>
<tr>
<th>$\delta$</th>
<th>Kernels</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RBF</td>
</tr>
<tr>
<td></td>
<td>d = 2</td>
</tr>
<tr>
<td>-0.2</td>
<td>73.7%</td>
</tr>
<tr>
<td>-0.3</td>
<td>68.4%</td>
</tr>
</tbody>
</table>

Table 2
Chapter 4

Conclusion

4.1: Analysis and Conclusion Of Results

The absolute least this study has shown is that Bollinger bands do affect the probability that options expire in or out-of-the-money and support vector machines can be used to analyze Bollinger bands and days to expiration to develop a good model that helps determine the probability an option will expire in or out-of-the-money. Based off the data in this study, with the exception of certain extreme situations, the original hypothesis has been proven to be effective using support vector machines. The original hypothesis for bullish markets is that call options should be bought and put options should be sold at times when the Bollinger bands for the stock are low. In bearish markets, put options should be bought when the Bollinger bands for the stock are high. The call options are not tested in a bearish market since the market moves so fast down in bearish markets.

According to the results of the tests, they are very supportive of the hypothesis in certain situations and the support vector machines have been proven to be very effective in certain situations. The first situation analyzed to use Bollinger bands with this strategy is in a bullish market. Generally, in bullish markets, calls should be bought and put options should be sold. The predictions for put options in bullish markets proved to be more accurate than the call options; however, there is a suspicion this is influenced by the fact that so many put options in a bullish market expire out-of-the-money. The models with the polynomial kernels for the call options with a value for $\delta$ being 0.2 are more supportive of the hypothesis and more accurate in their predictions than the RBF kernel.
The models are more accurate that have a value for $|\delta|$ being lower or the options being further out-of-the-money.

The bearish markets also performed as expected with the put options. The probabilities increased for the put options to expire in-the-money for higher %B values and decreased for lower %B values. Some tests are proven to be more accurate for put options in bearish markets than call options in bull markets. However, due to the market moving so quickly and being so volatile in bearish markets, the range of accuracy is very large which cautions using this strategy in bearish markets.

Although the Bollinger bands are not as effective in certain situations as in others, the support vector machines in this test have proven that Bollinger bands do affect the probability of the outcome of options nonetheless. The reason why is because for both the bullish and bearish markets, the support vector machines show that despite every option in this test having a value of 0.2 or 0.3 for $\delta$, in certain situations there is still more of a chance the option will expire in-the-money. The reason why this is significant is because the value $\delta$ for options is used to be a theoretical probability based off past models developed such as the Black-Scholes model that an option expires in-the-money. Theoretically, this means that every option in this study, theoretically should have a greater chance of expiring out-of-the-money than in-the-money. However, the support vector machines show that even though these options have a $\delta$ less than 0.5, in certain situations, the probability of the option expiring in-the-money is greater than the probability of it expiring out-of-the-money. In other words, even the options that have a $\delta$ of 0.2 which theoretically says the probability of the option expiring in-the-money is only 20%, the support vector machines show that in certain situations, these same options
actually have a greater than 50% chance of expiring in-the-money based off where the current price is in relation to the Bollinger bands. This proves the usefulness of support vector machines for predicting options as well as the significance of Bollinger bands. These observations with probability being affected by the Bollinger bands can allow an investor to see value in options other models can’t see including the Black Scholes Model.
References


11. [http://upload.wikimedia.org/wikipedia/commons/2/2a/Svm_max_sep_hyperplane_withmargin.png](http://upload.wikimedia.org/wikipedia/commons/2/2a/Svm_max_sep_hyperplane_withmargin.png)

12. [http://docs.opencv.org/_images/sample-errors-dist.png](http://docs.opencv.org/_images/sample-errors-dist.png)


