Stock Market Forecasting

Mohammed Zeb Alam
zebavailable@gmail.com
IT DEPARTMENT
BVCOE DELHI

Abstract

Forecasting the Stock Market is regarded as the most challenging task of a time series Forecasting. It is regarded as one of the most important forecasting of the century as it can generate us direct profits. Due to its complex nature an exact prediction model is difficult to establish. In this paper we predict the stock price movement of New York Stock Exchange (NYSE) based on SVM Classification Technique using the Kernel Functions like RBF, Polynomial, Sigmoid, Linear. We also strengthen the technique by proposing a Novel Kernel Function named Poly-Sigmoid Kernel. Given that Data is non linear and sparse it makes prediction tough. Our Model takes in forty-one-time variant features as the input for our model. It is used to predict next-day stock trend and suggest whether to buy or sell a Particular Stock. Experimental analysis of our prediction model shows that efficiency range from 53.4 % to 73.1% and 66% for our Poly Sigmoid Kernel Function. Investment returns range from 54.66% to 86.7% and 86.7 % for our Poly Sigmoid Kernel.

Introduction

The market is intricate with progressive data. Prediction of financial data is complex as it has high degree of uncertainty with many factors[7] . The market and stock prices can vary according to the social and economic factors which are unpredictable. The multi-dimensional data mixed with unstructured nature , high degree of uncertainty and noise makes it difficult for prediction. [8]

Support vector machine (SVM) is an algorithm that classifies the data using kernel functions. [9,20,21] It's Applications vary from classification of images to handwriting recognitions to pattern prediction. SVM can perform both nonlinear and linear classification efficiently using kernel functions. Kernel methods are used for pattern testing in SVM.

Kernel functions allows them to work in multi-dimensional space. In SVM classification model examples are treated as data points in space which belong to different categories. If data is labeled for learning, then supervised learning could be performed otherwise clustering is needed for unsupervised learning. The accuracy of prediction in most tests are depended on the divergence from expected values. General approaches for stock market prediction are either Univariate or multivariate analysis. In Multivariate Environment any input variable whether it is dependent on output or not can be incorporated as input variable while in Univariate analysis the input variable are restricted to time series being forecasted. ARIMA method is the example of Univariate Method. ANN (Artificial Neural Network) and SVM are used as multivariate method for the prediction of stock market as financial time series is non linear and non stationary . Multivariate Time series models depends on various indicators which could be fundamental , technical and inter market. The goal of our study is to find the accuracy of various kernels methods of SVM to find patterns in stock market prices and carry out a successful trade.In our study we combine these indicators to act as inputs. The Remainder of the paper is divided into four Sections. First Section being the Literature Review, Second Section provide us a little in-depth about the SVM, Third section describes the Research data , Forth Section provides the experimental results about the prediction model and a comparative analysis among various kernel and fifth section contains conclusion and future scope

Literature Review

In the past various stock market analysis were based on time series multivariate analysis . In the recent time various machine learning methodologies like SVM , Neural Network , Genetic Algorithm , Naïve Bayes have been Applied to predict the movement of stock market In this study we are going to use SVM and various kernels of SVM to predict make a prediction model to predict the movement of NYSE . In [22] , [23] we study the movement of US , UK (London Stock Exchange) respectively using the principle of ANN(Artificial Neural Network)for the stock prediction . In [24] we study the making of the prediction model of Stock market using the concept of SVM . In [25] Chen et all , Applies SVM and Artificial Neural Network using back propagation methodology for the study of six Asian stock markets . in [26] bankruptcy prediction using Support Vector machine is done and a analysis with Neural Network is also carried out . In , Tay et al. [27] have used Chicago mercantile markets data sets to see the feasibility for financial time series forecasting by applying the method of stock market prediction .

Support vector machine

SVM is a supervised learning technique applied to analyze our time variant stock market data. With a set of training examples, belonging to either category. SVM training algorithm is used to build a technique to place new examples to either category. The examples are mapped into the categories .It is used to create a boundary between categories as wide as possible. New data is mapped and placed in the specified category for labeled data. For data which is not labeled unsupervised learning is used, which attempts to find a natural clustering of the data to groups. Thus, new data is mapped to these formed groups. The clustering algorithm which provides an improvement to the support vector machines is called **support vector clustering [2].**

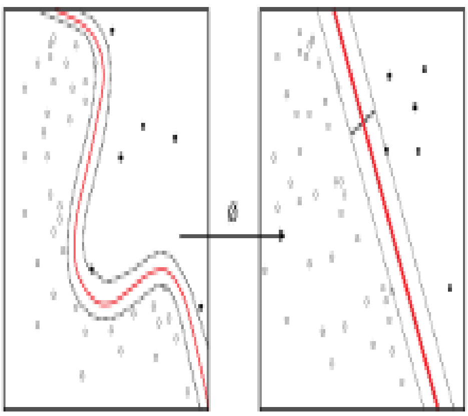


Fig 1. Non Linear Mapping to Linear Mapping by applying Kernel Function

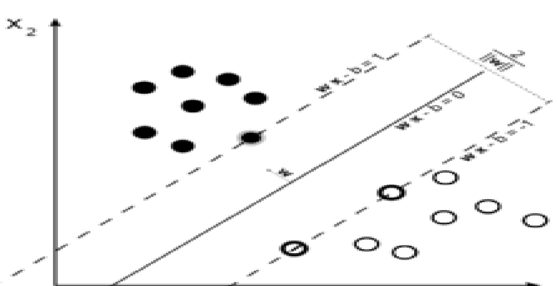
Mathematics for SVM

Training dataset of n points is of the following form

(vec(x,1),(y,1)),.......,(vec(x ,n) , (y,n))

where the y(i) are either 1 or −1, each indicating the class to which the point vec(x(i)) belongs. Each vec(x(i)) is a p-dimensional real vector. We want to find the "maximum-margin hyper plane" that divides the group of points vec(x(i)) for which y(i)=1 from the group of points for which y(i) = -1, which is defined so that the distance between the hyperplane and the nearest point vec(x(i)) from either group is maximized.[1]

A Hyper plane written as the set of points vec(x(i))[1]



vec(w).vec(x)+b=0,

where,

vec(w) is the normal vector to the plane.

b/||vec(w)|| is offset of the hyper plane.

Broadly two Types of Margins are there :
-Hard Margin.
-Soft Margin

Nonlinear classification

For the task of Non Linear Classification Kernel Trick is used . Kernel trick is used to widen the gap between the two categories. The algorithm is there, so that, dot product is replaced by a nonlinear kernel function. This allows the algorithm to fit a transformed feature space with maximum-margin hyper plane.
Some kernels:
Polynomial (Homogeneous):

$$k(\vec{x_i}, \vec{x_j}) = (\vec{x_i} \cdot \vec{x_j})^d$$

Polynomial (inhomogeneous)

$$k(\vec{x_i}, \vec{x_j}) = (\vec{x_i} \cdot \vec{x_j} + 1)^d$$

Gaussian radial basis function:

$$k(\vec{x_i}, \vec{x_j}) = \exp(-\gamma \|\vec{x_i} - \vec{x_j}\|^2)$$

for,

Hyperbolic tangent:
for some, k>0 and c<0

$$k(\vec{x_i}, \vec{x_j}) = \tanh(\kappa \vec{x_i} \cdot \vec{x_j} + c)$$

Computing the SVM classifier

$$\begin{aligned} b = \vec{w} \cdot \varphi(\vec{x}_i) - y_i &= \left[\sum_{k=1}^{n} c_k y_k \varphi(\vec{x}_k) \cdot \varphi(\vec{x}_i)\right] - y_i \\ &= \left[\sum_{k=1}^{n} c_k y_k k(\vec{x}_k, \vec{x}_i)\right] - y_i. \end{aligned}$$

SVM classifier amounts to minimizing an expression of the form $\left[\frac{1}{n}\sum_{i=1}^{n} \max\left(0, 1 - y_i(w \cdot x_i + b)\right)\right] + \lambda \|w\|^2. \qquad (2)$

Kernel trick

Kernel is essentially a mapping function - one that transforms a given space into some other (usually very high dimensional) space.

With a kernel function k such that:

$$k(\vec{x}_i, \vec{x}_j) = \varphi(\vec{x}_i) \cdot \varphi(\vec{x}_j)$$

vec(w) in the transformed space satisfies

$$\vec{w} = \sum_{i=1}^{n} c_i y_i \varphi(\vec{x}_i),$$

Where ,c(i) is obtained through optimization of the following[1]

$$\begin{aligned} \text{maximize } f(c_1 \ldots c_n) &= \sum_{i=1}^{n} c_i - \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} y_i c_i (\varphi(\vec{x}_i) \cdot \varphi(\vec{x}_j)) y_j c_j \\ &= \sum_{i=1}^{n} c_i - \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} y_i c_i k(\vec{x}_i, \vec{x}_j) y_j c_j \end{aligned}$$

we can find some index ***i*** such that $0 < c_i < (2n\lambda)^{-1}$, so that $\varphi(\vec{x}_i)$ lies on the boundary of the margin in the transformed space, and then solve

$$\vec{z} \mapsto \operatorname{sgn}(\vec{w} \cdot \varphi(\vec{z}) + b) = \operatorname{sgn}\left(\left[\sum_{i=1}^{n} c_i y_i k(\vec{x}_i, \vec{z})\right] + b\right).$$

New points can be classified by computing

$$\gamma = \frac{1}{2\sigma^2}$$

Kernel method

Kernel methods are a class of algorithms for pattern analysis with SVM. Pattern analysis finds and study relations in datasets. Kernel methods only require user specified kernel over a pair of data points in a raw representation.

Some of the application of kernel methods are Geo Statistics, Kinging, inverse distance weighting, 3D reconstruction, bioinformatics, chemo informatics, information extraction and handwriting recognition.

Polynomial kernel

It is a kernel function commonly used with SVM. It devices features of its input samples. It also uses a bunch of combinations of these features for classification..

For degree-*d* polynomials, the polynomial kernel is defined as [2']

$$K(x, y) = (x^{\mathsf{T}} y + c)^d$$

where *x* and *y* are vectors in the input space and c>=0 is a free parameter trading off the influence of higher-order versus lower-order terms in the polynomial [5] . When *c* = 0, the kernel is called homogeneous. [3]

K corresponds to an inner product in a feature space based on some mapping φ

$$K(x, y) = \langle \varphi(x), \varphi(y) \rangle$$

Radial basis function kernel

With two samples x and x', given as features in some outer space, is defined as [4]

$$K(\mathbf{x}, \mathbf{x}') = \exp\left(-\frac{||\mathbf{x} - \mathbf{x}'||^2}{2\sigma^2}\right)$$

$||\mathbf{x} - \mathbf{x}'||^2$ = squared Euclidean distance between the two feature vectors. An equivalent, but simpler, definition involves a parameter .RBF kernel decreases with distance and ranges between zero and one. Its feature space has infinite number of dimensions .[6]

Sigmoid Kernel

With two samples x and x’ , given as features in some outer space , is defined as :

$$\tanh(\gamma\langle x, x'\rangle + r)$$

the gamma parameter defines how far the influence of a single training example reaches, with low values meaning ‘far’ and high values meaning ‘close’. The gamma parameters can be seen as the inverse of the radius of influence of samples selected by the model as support vectors.

Linear Kernel

With two samples x and x’ , given as features in some outer space , is defined as :

$$\langle x, x' \rangle.$$

Poly Sigmoid Kernel

This is our custom proposed Kernel which is a linear combination of polynomial and sigmoid kernel given as:

$$(x^{T}y+c)^{d} \quad + \tanh(\gamma\langle x, x'\rangle + r)$$

3. Research Data and Experimentation

3.1 Research Data

The Data used in the study comprises of the fundamental and technical indicators of the S&P 500 and New York stock exchange (NYSE)[10] . We have picked 560 Stocks over the time span of 2003-01-04 to 2013 – 05-16. Our Aim here is to perform the basic analysis and based on the features and previous trends predict whether the investment will be profitable or not. Our Data Set Consists of 41 features as shown in table 1:

Sno	Feature	Indicate
1	Date	Indicate the date and time
2	Ticker	Represents the acronym for the stock
3	Price	Stock Price
4	stock_p_change	%change in stock as compared to previous value
5	SP500_p_change	%change in price of SP_500 as compared to previous value
6	Difference	stock_p_change-sp500_p_change
7	DE Ratio	Total Debt/Equity
8	Trailing P/E	Trailing P/E (ttm, intraday)
9	Price/Sales	Price/Sales (ttm)
10	Price/Book	Price/Book(ttm)
11	Profit Margin	Profit Margin(ttm)
12	Operating Margin	Operating Margin(ttm)
13	Return on Assets	Return on Assets
14	Return on Equity	Return on Equity
15	Revenue Per Share	Revenue Per Share
16	Market Cap	Market Cap
17	Enterprise Value	Enterprise Value
18	Forward P/E	Forward P/E (fye 31-Dec-05)
19	Enterprise Value/Revenue	Enterprise Value/Revenue
20	Enterprise Value/EBITDA	Enterprise Value/EBITDA
21	Revenue	Revenue
22	Gross Profit	Gross Profit
23	EBITDA	Earnings before interest, tax, depreciation and amortization
24	Net Income Avl to Common	Net Income Avl to Common
25	Diluted EPS	Diluted EPS
26	Earnings Growth	Earnings Growth
27	Revenue Growth	Revenue Growth
28	Total Cash	Total Cash
29	Total Cash Per Share	Total Cash Per Share
30	Total Debt	Total Debt
31	Current Ratio	Current Ratio
32	Book Value Per Share	Book Value Per Share
33	Cash Flow	Cash Flow
34	Beta	Beta
35	Held by Insiders	Shares Held by Insiders
36	Held by Institutions	Shares Held by Institutions
37	Shares Short Today	Shares Short Today
38	Short Ratio	Short Ratio
39	Short % of Float	Short % of Float
40	Shares Short Prior	Shares Short Prior
41	Status	Status

Table 1. Data Set

The Data was collected in the form of HTML Files from Yahoo Finance [11] as shown in Fig 1 . We converted the Data from unstructured Format (HTML) to Structured Format (Comma Separated Value Format) as shown in Fig 2 .

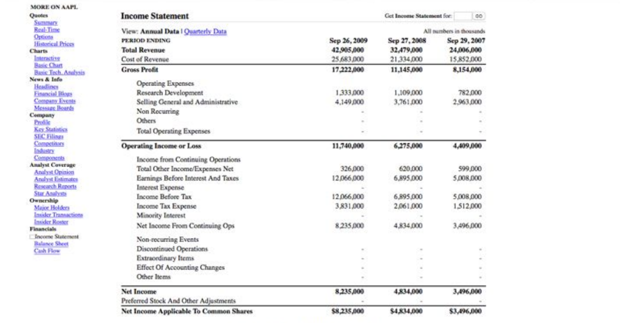
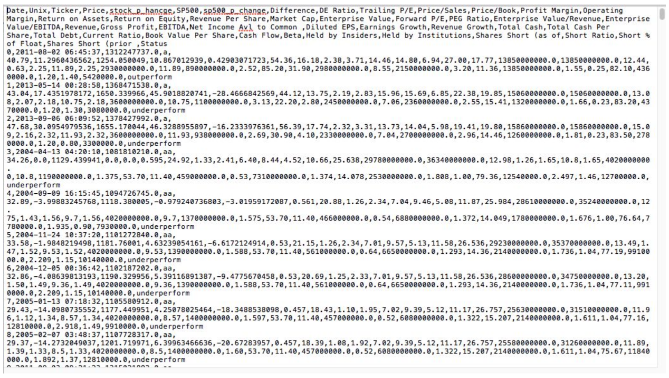


Fig 1 HTML Document



```
Date,Unix,Ticker,Price,stock_p_hancge,SP500,sp500_p_change,Difference,DE Ratio,Trailing P/E,Price/Sales,Price/Book,Profit Margin,Operating
Margin,Return on Assets,Return on Equity,Revenue Per Share,Market Cap,Enterprise Value,Forward P/E,PEG Ratio,Enterprise Value/Revenue,Enterprise
Value/EBITDA,Revenue,Gross Profit,EBITDA,Net Income Avl to Common ,Diluted EPS,Earnings Growth,Revenue Growth,Total Cash,Total Cash Per
Share,Total Debt,Current Ratio,Book Value Per Share,Cash Flow,Beta,Held by Insiders,Held by Institutions,Shares Short (as of,Short Ratio,Short %
of Float,Shares Short (prior ,Status
0,2011-08-02 06:45:37,1312247737.0,a,
40.79,11.2960436562,1254.050049,10.867012939,0.42903071723,54.36,16.18,2.38,3.71,14.46,14.80,6.94,27.00,17.77,13850000000.0,13850000000.0,12.44,
0.63,2.25,11.89,2.25,2930000000.0,11.89,890000000.0,2.52,85.20,31.90,2980000000.0,8.55,2150000000.0,3.20,11.36,13850000000.0,1.55,0.25,82.10,436
0000.0,1.20,1.40,5420000.0,outperform
1,2013-05-14 00:28:58,1368471538.0,a,
43.04,17.4351978172,1650.339966,45.9018820741,-28.4666842569,44.12,13.75,2.19,2.83,15.96,15.69,6.85,22.38,19.85,15060000000.0,15060000000.0,13.0
8,2.07,2.18,10.75,2.18,3600000000.0,10.75,1100000000.0,3.13,22.20,2.80,2450000000.0,7.06,2360000000.0,2.55,15.41,1320000000.0,1.66,0.23,83.20,43
70000.0,1.20,1.30,3080000.0,underperform
2,2013-09-06 06:09:52,1378427992.0,a,
47.68,30.0954979536,1655.170044,46.3288955897,-16.2333976361,56.39,17.74,2.32,3.31,13.73,14.04,5.98,19.41,19.80,15860000000.0,15860000000.0,15.0
9,2.16,2.32,11.93,2.32,3600000000.0,11.93,938000000.0,2.69,30.90,4.10,2330000000.0,7.04,2700000000.0,2.96,14.46,1260000000.0,1.81,0.23,83.50,278
0000.0,1.20,0.80,3300000.0,underperform
3,2004-04-13 04:20:10,1081810210.0,aa,
34.26,0.0,1129.439941,0.0,0.0,0.595,24.92,1.33,2.41,6.40,8.44,4.52,10.66,25.638,29780000000.0,36340000000.0,12.98,1.26,1.65,10.8,1.65,4020000000
.
0,10.8,1190000000.0,1.375,53.70,11.40,459000000.0,0.53,7310000000.0,1.374,14.078,2530000000.0,1.808,1.00,79.36,12540000.0,2.497,1.46,12700000.0,
underperform
4,2004-09-09 16:15:45,1094726745.0,aa,
32.89,-3.99883245768,1118.380005,-0.979240736803,-3.01959172087,0.561,20.88,1.26,2.34,7.04,9.46,5.08,11.87,25.984,28610000000.0,35240000000.0,12
.
75,1.43,1.56,9.7,1.56,4020000000.0,9.7,1370000000.0,1.575,53.70,11.40,466000000.0,0.54,6880000000.0,1.372,14.049,1780000000.0,1.676,1.00,76.64,7
780000.0,1.935,0.90,7930000.0,underperform
5,2004-11-24 10:37:20,1101272840.0,aa,
33.58,-1.9848219498,1181.76001,4.63239054161,-6.6172124914,0.53,21.15,1.26,2.34,7.01,9.57,5.13,11.58,26.536,29230000000.0,35370000000.0,13.49,1.
47,1.52,9.53,1.52,4020000000.0,9.53,1390000000.0,1.588,53.70,11.40,561000000.0,0.64,6650000000.0,1.293,14.36,2140000000.0,1.736,1.04,77.19,99100
00.0,2.209,1.15,10140000.0,underperform
6,2004-12-05 00:36:42,1102187202.0,aa,
32.86,-4.08639813193,1190.329956,5.39116891387,-9.4775670458,0.53,20.69,1.25,2.33,7.01,9.57,5.13,11.58,26.536,28600000000.0,34750000000.0,13.20,
1.50,1.49,9.36,1.49,4020000000.0,9.36,1390000000.0,1.588,53.70,11.40,561000000.0,0.64,6650000000.0,1.293,14.36,2140000000.0,1.736,1.04,77.11,991
0000.0,2.209,1.15,10140000.0,underperform
7,2005-01-13 07:18:32,1105580912.0,aa,
29.43,-14.0980735552,1177.449951,4.25078025464,-18.3488538098,0.457,18.43,1.10,1.95,7.02,9.39,5.12,11.17,26.757,25630000000.0,31510000000.0,11.9
6,1.12,1.34,8.57,1.34,4020000000.0,8.57,1400000000.0,1.597,53.70,11.40,457000000.0,0.52,6080000000.0,1.322,15.207,2140000000.0,1.611,1.04,77.16,
12810000.0,2.918,1.49,9910000.0,underperform
8,2005-02-07 03:48:37,1107728317.0,aa,
29.37,-14.273204903,1201.719971,6.39963466636,-20.67283957,0.457,18.39,1.08,1.92,7.02,9.39,5.12,11.17,26.757,25580000000.0,31260000000.0,11.89,
1.39,1.33,8.5,1.33,4020000000.0,8.5,1400000000.0,1.60,53.70,11.40,457000000.0,0.52,6080000000.0,1.322,15.207,2140000000.0,1.611,1.04,75.67,11840
000.0,1.892,1.37,12810000.0,underperform
```


Fig 2 CSV (Comma Separated Value)

Here Status feature is the feature which indicate whether a Stock Price outperforms or underperforms the SP 500 Index for the same time duration. It acts as the dependent variable or Class Variable [12] for our classifier. All other features act as the independent variables [12] for our classifiers to check the variation of dependent variable based on independent variable .

3.2 Preprocessing

Next Step in Data Analysis is Data Preprocessing . Data Processing means to deal with missing values, finding outliers (noise) and inconsistencies in the data set. Here as we are dealing with Stock Market data set we won't be appropriate to replace the N/A or inconsistent data with mean values as suggested by scientific community [13] . So we delete the inconsistent and N/A Data. It has been proven that SVM performs better when data has zero mean and unit variance [14,15,16]. So , We Standardize the features by removing the mean and scaling to unit variance for better analysis.

3.3 Preparation of the class variable for our Prediction Model

To prepare our data for the task of classification we have to prepare the class variable for the task of classification is the most important step [17]. In our case this class variable is status which can be subdivided into two categories outperform (1) or underperform (0). Status class is prepared by using the percent change for value of the stock (stock_p_change) and percent change for the value of SP500 index(SP500_p_change) for the same duration. Status is given an outperform value if stock_p_change > SP500_p_change and underperforms if stock_p_change <= SP500_p_change . If Status is categorized into Outperform then we should buy a particular stock. If Status is categorized into Underperform, then we should sell a particular Stock

3.4 Prediction Model

Our main objective to find the value of output variable (Status) based on set of input features (All features except price, stock_p_change , SP500_p_change , difference) .

3.5 Comparisons Parameters

Accuracy: It is defined as the number of true predictions divided by total number of predictions [28]
Precision: It is defined as number of true positives(tp) divided by sum of true positives and true negatives(tn)[28]
Precision = tp/(tp+tn)
Recall: It is defined as number of true positives (tp) divided by sum of true positives(tp) and false negatives (fn)[28]
recall = tp/(tp+fn)

TP= true positive, TN=true negative, FP=false positive, FN=false negative

	Prediction of model: positive	Prediction of model: negative
Truth: positive	TP	FN
Truth: negative	FP	TN

Here :
tp – True Positives is used to tell the number of Positives Predictions which are actually positives
tn – True Negatives is the number of Negative Predictions which are actually negatives
fp – False Positives is the number of Positive Predictions which are actually negatives
fn – False Negatives is the number of Negative Predictions which are actually positives.

F Score = 2* precision * recall /(precision +recall)[28]

4. Experiments Results

In this paper stock market forecast is done by the classification techniques SVM with Linear Kernel, SVM with Polynomial Kernel of degree four, SVM with RBF Kernel, SVM with Sigmoid Kernel and SVM with Poly Sigmoid Kernel . Data Set of 2991 records of observations is separated into training and testing sets for the task of classification. We execute the process of Monte Carlo cross-validation [18] where 70 % was our training data and 30 % was our testing data. We test the classifiers based on its efficiency [19], time taken to train the classifier , precision[19] , recall[19] and F score[19] .

Algorithm	Efficiency	Time Taken (in sec)
SVM with Linear Kernel	65.2 %	0.453
SVM with Polynomial Kernel of degree four	58.5 %	0.350

SVM with RBF Kernel	73.1 %	0.325
SVM with Sigmoid Kernel	53.4 %	0.306
SVM with Poly Sigmoid Kernel	66 %	1.09

Table 2 . Efficiency and time taken by the algorithms

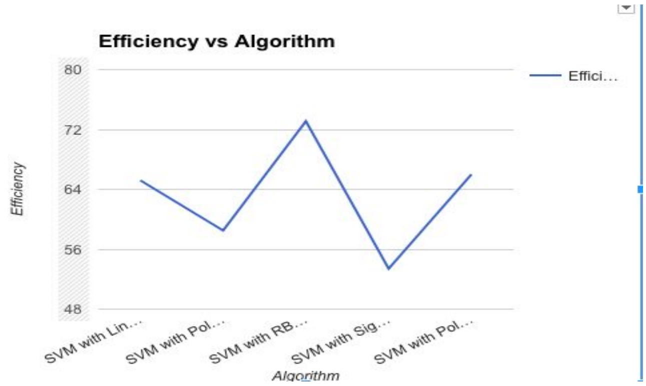


Fig 2.1. Variation of Efficiency with Algorithm (Line Graph)

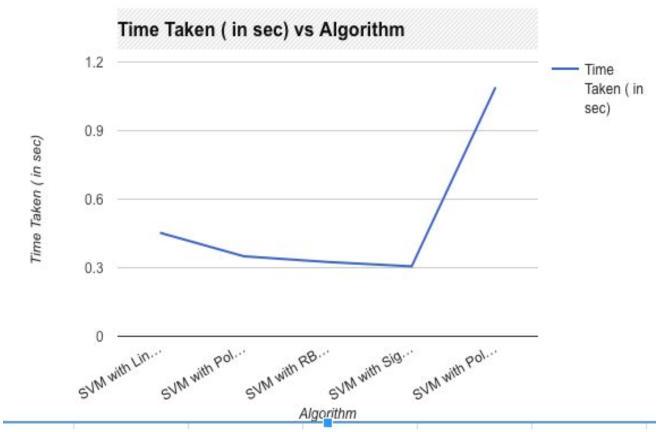


Fig 2.2 Time Taken Vs Algorithm

According to our analysis we can clearly see SVM with Linear (65.2%), RBF Kernel (73.1 %) and Poly Sigmoid Kernel (66 %) tend to perform better Polynomial (58.5 %) and Sigmoid Kernel (53.4%). When we take time in aspect Sigmoid take the least time and Poly Sigmoid take the most.

Algorithm	Precision	Recall	F - Score
SVM with Linear Kernel	0.612	0.810	0.697
SVM with Polynomial Kernel of degree four	0.549	0.991	0.706
SVM with RBF Kernel	0.686	0.801	0.739
SVM with Sigmoid Kernel	0.494	1.0	0.661
SVM with Poly Sigmoid Kernel	0.82	0.63	0.71

Table 3 . Precision , Recall and F Score

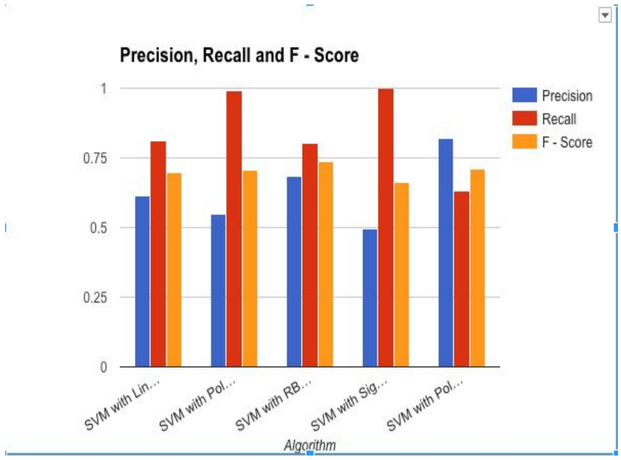
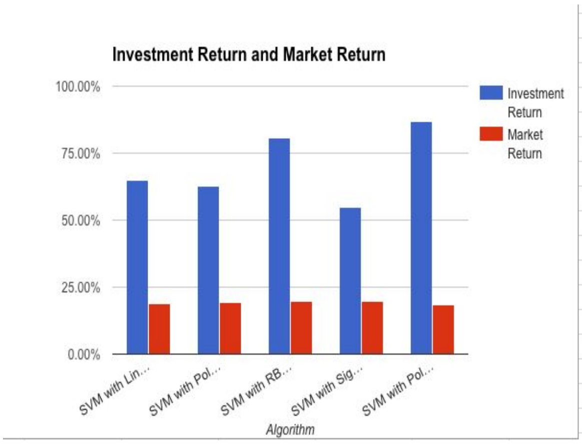


Fig 3 Precision , Recall and F-Score V/s Algorithm

Here we could clearly observe RBF Kernel has the best f score relative to others. To Demonstrate the efficiency of our Model we simulated a virtual trade of 10,000 $ on the same duration as the test set.

Algorithm	Investment Return	Market Return	Trades Done
SVM with Linear Kernel	65.05%	19.07%	578
SVM with Polynomial Kernel of degree four	62.87%	19.22%	814
SVM with RBF Kernel	80.89%	19.64%	523
SVM with Sigmoid Kernel	54.66%	19.62%	896
SVM with Poly Sigmoid Kernel	86.7%	18.6%	587



It shows that RBF and Poly Sigmoid Kernel Gives the maximum return in minimum number of trades thus saving us on trading costs as well followed by linear, polynomial and sigmoid kernel

5. Conclusion and Future Scope

In this paper we study the algorithm named SVM for forecasting and prediction of stock market. It turned out to be an amazing tool for the same. Here we also propose a novel kernel called a Poly Sigmoid Kernel which is an amalgamation of Polynomial and Sigmoid Kernel .We then indulge into analysis of various kernels used with SVM and found out that RBF is the most promising Kernel in terms of Efficiency, F Score . Our Novel Kernel Poly Sigmoid Kernel gives the maximum return on a sample trade on the contrary it takes the maximum time which shows a potential with a hiccup . Linear Kernel which seem promising in terms of efficiency and return but takes the most time to train. Polynomial of degree four is an average Kernel function for this task . Sigmoid Kernel is most performing kernel which should be avoided for stock prediction task . Each Kernel Functions has their own advantages and disadvantages and may find different applications in different fields. We Believe weakness of one kernel can be balanced by the strengths of another kernel function or another classification function . So , In Near Future we would work on combing different classifications techniques and comparing it to preexisting ones like SVM , Naïve Bayes etc.

6. References:

[1]Andrew,A.M(2000).An Introduction to Support Vector Machines and other Kernel Based Learning Methods By Nello Christianni and John Shawe-Taylor , Cambridge University Press , Cambridge 2000, xiii+,189 pp , ISBN 0-521-78019-5 .

[2]Ben-Hur, Asa, Horn, David, Siegelmann, Hava, and Vapnik, Vladimir; "Support vector clustering" (2001) Journal of Machine Learning Research, 2: 125–137.

[3] Goldberg, Y., & Elhadad, M. (2008, June). splitSVM: fast, space-efficient, non-heuristic, polynomial kernel computation for NLP applications. In *Proceedings of the 46th Annual Meeting of the Association for Computational Linguistics on Human Language Technologies: Short Papers* (pp. 237-240). Association for Computational Linguistics.

[4]Shashua, Amnon (2009). "Introduction to Machine Learning: Class Notes 67577". arXiv:0904.3664v1 [cs.LG].

[5]Jean-Philippe Vert, Koji Tsuda, and Bernhard Schölkopf (2004). "A primer on kernel methods". Kernel Methods in Computational Biology.

[6] Schölkopf, Bernhard, and Alexander J. Smola. *Learning with kernel regularization, optimization, and beyond*. MIT press, 2002. APA

[7] Musavi, M. T., Ahmed, W., Chan, K. H., Faris, K. B., & Hummels, D. M. (1992). On the training of radial basis function classifiers. *Neural networks*, *5*(4), 595-603.

[8] Abu-Mostafa YS, Atiya AF. Introduction to financial forecasting. Applied Intelligence 1996;6:205–13.

[9] Hall JW. Adaptive selection of US stocks with neural nets. In: Deboeck GJ, editor. Trading on the edge: neural, genetic, and fuzzy systems for chaotic financial markets. New York: Wiley; 1994. p. 45–65.

[10] Vapnik VN. Statistical learning theory. New York: Wiley; 1998.

[11] Elliott, William B., and Richard S. Warr. "Price pressure on the NYSE and NASDAQ: Evidence from S&P 500 index changes." *Financial Management* (2003): 85-99.

[12] https://finance.yahoo.com

[13] Lieberman, Norman P. "Dependent and Independent Variables." *Troubleshooting Process Plant Control*: 21-28.

[14] Mitra, Sanjit Kumar, and Yonghong Kuo. *Digital signal processing: a computer-based approach*. Vol. 2. New York: McGraw-Hill, 2006.

[15] Forman, G., Scholz, M., & Rajaram, S. (2009, June). Feature shaping for linear SVM classifiers. In *Proceedings of the 15th ACM SIGKDD international*

conference on Knowledge discovery and data mining (pp. 299-308). ACM.

[16] Laanaya, H., Abdallah, F., Snoussi, H., & Richard, C. (2011). Learning general Gaussian kernel hyperparameters of SVMs using optimization on symmetric positive-definite matrices manifold. *Pattern Recognition Letters*, *32*(13), 1511-1515.

[17] Chapelle, O., & Keerthi, S. S. (2008, August). Multi-class feature selection with support vector machines. In *Proceedings of the American statistical association*.

[18] Berson, Alex, and Stephen J. Smith. *Data warehousing, data mining, and OLAP*. McGraw-Hill, Inc., 1997.

[19] Xu, Q. S., & Liang, Y. Z. (2001). Monte Carlo cross validation. *Chemometrics and Intelligent Laboratory Systems*, *56*(1), 1-11.

[20] Powers, David Martin. "Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation." (2011).

[21] Vapnik VN. An overview of statistical learning theory. IEEE Transactions of Neural Networks 1999;10:988–99.

[22]Cristianini N, Taylor JS. An introduction to support vector machines and other kernel-based learning methods. New York: Cambridge University Press; 2000.

[23] Refenes, A-P., Zapranis, A.D. and Francis, G. (1995) „Modeling stock returns in the framework of APT: a comparative study with regression models‟, Neural Networks in the Capital Markets, pp.101–125.

[24] Tsibouris, G. and Zeidenberg, M. (1995) „Testing the efficient markets hypothesis with gradient descent algorithms‟, Neural Networks in the Capital Markets, pp.127–136.

[25] Tay FEH, Cao LJ., 2001, Improved financial time series forecasting by combining support vector machines with self- organizing feature map. Intelligent Data Analysis.

[26] Wun-Hua Chen and Jen Ying Shih., "Comparison of support vector machines and back propagation neural networks in forecasting the six major Asian stock markets", International Journals Electronics Finance, Vol. 1, 2006, pp.49-67.

[27] Chih Fong Tsai., "Financial decision support using neural network and support vector machines", Expert Systems, Vol. 25, 2008, pp.380-393.

[28] Tay, F.E.H. and Cao, L., "Application of support vector machines in financial time-series forecasting", Omega, Vol. 29, 2001, pp.309–317.

[29] Goutte, Cyril, and Eric Gaussier. "A probabilistic interpretation of precision, recall and F-score, with implication for evaluation." *European Conference on Information Retrieval*. Springer Berlin Heidelberg, 2005.