

STOCK TREND PREDICTION MODEL BASED ON RBF-SVM ALGORITHM**¹Shalinigayathri. D and ²A.S. Arunachalam**¹Research Scholar, Department of Computer Science, School of Computing Sciences, Vels Institute of Science, Technology and Advanced Studies (VISTAS), Tamilnadu India.²Associate Professor, Department of Computer Science, School of Computing Sciences, Vels Institute of Science, Technology and Advanced Studies (VISTAS), Tamilnadu India¹san.thiru@gmail.com and ²arunachalam1976@gmail.com**ABSTRACT**

In current times, the pandemic situation of covid-19 has made the stock market more volatile. Even though people should not stop the research and prediction on the stock market. With the help of recent Machine learning technologies, the stock market data analysis is completely reliable and efficient. The Machine learning model provides the opportunity for traders to choose more profitable stocks. The proposed work focus to improve the prediction stock market by the network model based on SVM. The forecasting technique is analyzed with SVM kernel models like RBF, Polynomial, and Linear. Even though the analysis of various kernel models of SVM, the model works the best when the RBF kernel model is used and shows the best accuracy and prediction of stock price.

Index Terms- Prediction, Classification, Regression, Optimization, Propagation, Volatile.

I. INTRODUCTION

A Financial market gives a better way of investment of financial resources in the stock market leads to buying and selling the values confidentially. At presently trading our financial resources in an organized structure provides great opening with high profit. The financial market can be categorized by financial claim and dept. instrument. The financial claim can trade in the stock market either a fixed amount or a residual amount. The debt instrument is one of the financial assets which can trade as a debt market. Financial assets are treated as equity instruments. An equity claim sanction for the investors to receive the fixed amount for their preferred stock. In the fixed income market permits the classification of the debt instruments and preferred stock. The common stock market does not include preferred stock. A financial market is two types based on maturity claim as money market and capital market. Short-term debt instrument in a financial market is referred to as money market whereas long-term instruments are called the capital market. According to the maturity level, markets are classified as short term and long term. If the maturity level is less than one year, it's called short term and they are belong to the Money market. If the maturity level is more than one year, it's called long-term and they are belong to the Capital market. Equity instruments are classified into part of the capital market. They have no maturity date. Nowadays Machine learning techniques have more popular for time series analysis and stock market prediction. Nowadays, most traders choose an intelligent trading system to help predict prices. Stock- related data are very much huge, so machine learning algorithms are used to predict stock values.

This study proposed various SVM kernel models to predict stock market behaviour and related works also analysed various stock market prediction models. In this paper, various SVM kernel model was trained on NSE data and at last the result shows that SVM-RBF kernel gives a better result.

II. RELATED WORK

In [1] author has used FFNN and ARIMA model to analyze the performance of the stock over 35 Bangladeshi stocks and find that the ARIMA model gives Maximum prediction accuracy and FFNN show that the best algorithm in forecasting stock index and his study return that FFNN gives Maximum accuracy in the out of 35 stocks. Sayavong et al.[2] has also used the CNN model on Thai stock market and found the best prediction for 3 stocks such as BBL, CAPLL and PTT. The results indicate that the CNN model can actively identify the changing trend of stock price. Zixuan[3] forecasts SVM kernel models to realize the accurate stock prediction and

results show that RBF-SVM gives 65.64% accuracy, PCA-SVM gives 62.17% accuracy, GA-SVM gives 64.67%. Fang Liu[4] find the combination of LSTM and attention model applied on cluster stocks and the results show that LSTM model shows 90% accuracy in the stock price prediction while the combination of LSTM-attention model shows 91% accuracy in the same stock price. Rajni Nikhil and Nitin[5] have used Decision Tree Regressor model, Random forest Regressor and Support vector regressor model. They find the common attributes such as date of stock, open prices of the stock, etc., along with Covid-19 features such as New cases, Cumulative Cases, New death, Cumulative death of the day, and the results show that all model performs well when the Covid-19 features were included. Hemil N.Shah[6] has used the Levenberg Maruardt method on stock price which is part of the Neural Network model. It gives not only perfect accuracy and also takes less time and less memory usage. Sahil Vazirani and Abhishek et al[7] has used linear regression append with another linear regression model on the stock data and the result shows that minimal error was occurred and performed more efficiently, accurately than stand-alone model such as KNN, SVM. Gourav Bathla[8] has proposed an LSTM model on the stock price and gets better accuracy than SVR Model.

III. METHODOLOGY

A. SVM-Linear

SVM's purpose is to fragments the datapoints into two different classes by the separation of hyperplane. The values close of its are referred as support vectors. The plane which is divided the data points as different classes while the margin is the gap between the data points between two different classes. Here choosing the appropriate hyperplane it represents an efficient way of data separation. It can be used to connect any two observations as a dot product. The linear kernel gives faster results than the kernel functions which is best suited for text classifications. The linear kernel formula, whereas the product of value is r & r_i , is the summation of each input values are shown in the formula.

$$k(r, r_i) = \text{sum}(r \times r_i)$$

B. SVM-Polynomial

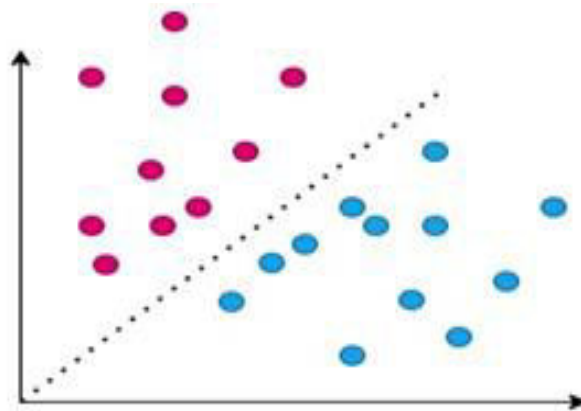


Fig 1. Linear separable data

The usage of polynomial kernel in SVM provides efficient transformation form linear inseparable data points in to discrete one. The technique of kernel promotes the input dimensional value from lower to higher space for optimized separation. By adding more dimensions to non-separable issues, the kernel transforms them into separable problems, thus it improves the accuracy and speed and reliability. The polynomial kernel springs key way for scatter and jumbled datapoint to desired format. It's a more generalized linear kernel that can distinguish between curved and nonlinear input spaces, the formula for polynomial kernel, whereas degree is denoted by t .

$$k(R, R_i) = 1 + \text{Sum}(R, R_i) \wedge t$$

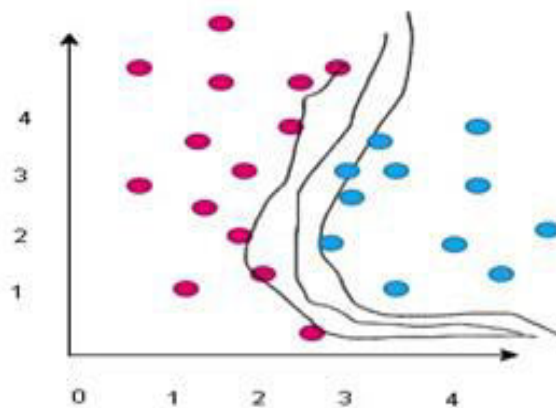


Fig 2. Polynomial SVM

C. SVM-Radial Basis Function

The usage of RBF kernel promotes good results for large number of data points which are unspecified values with greater accuracy. For Radial Basis Function is most suited for if the data points are either inseparable or non-linear. According to the gamma value, the model reaches its boundaries of anticipated one where the lower value indicates lower accuracy like linear model results, higher value of gamma represents overfitting and no regulation, intermediate gamma values give the optimized result accuracy. The misclassification is controlled by the regularization parameter which gives tolerance and robustness. The model which is best suited for if there is no need for prior knowledge data points, the formula is

$$K(r, r) = \exp(-\text{gamma} * \text{sum}(r - r_i, 2))$$

The value of Gamma range between the value of 0 to 1, whereas its default value is 0,1 is the excellent starting point.

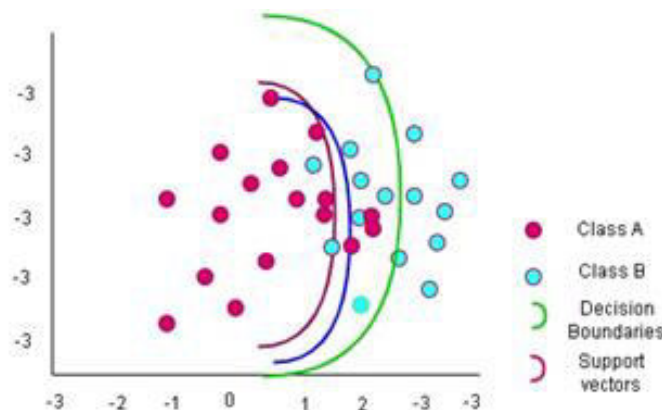


Fig 3. SVM-RBF

IV. STOCK PRICE PREDICTION BASED ON SVM

At first, Raw data is collected, then by using data preprocessing method unwanted, missed data are cleaned. Now the Data are divided into two categories. One set of data is going to train the model and another part of data are given to test the model. At the last, by using performance measure. The result will be predicted.

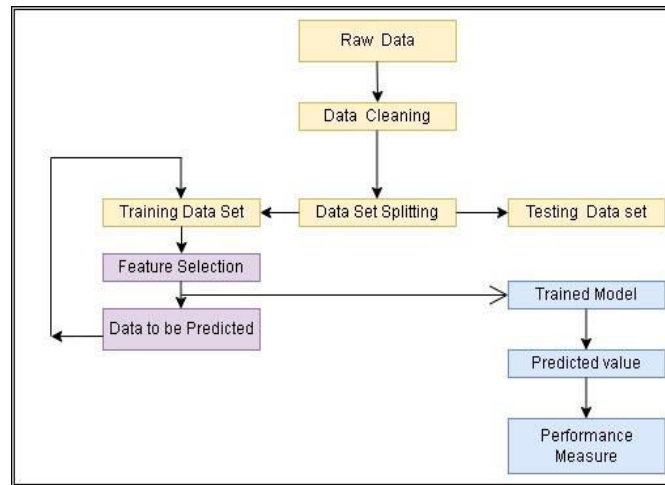


Fig 4. Flow chart of proposed methodology

A. Stock price prediction based on SVM: Data set establishment

In this paper, the experimental data which is taken from the NSE India website. TataElxsi, MindTree consider as a dataset. The dataset have 20 rows and 8 columns. A single day represented by a single row. The Experimental data for TataElxsi are:

Table 1: Experimental data for TataElxsi

Date	Open	High	Low	Close	Adj Close	Volume
03-01-22	5888	5983.75	5868	5893.65	5893.65	120686
04-01-22	5950	5989.5	5895.05	5977.75	5977.75	107717
05-01-22	6000	6027	5884.95	5963.8	5963.8	106351
06-01-22	5873	5901.75	5805	5873.55	5873.55	93465
07-01-22	5900	5938.85	5782	5799.55	5799.55	112039
10-01-22	5861.05	6020	5800	5992.95	5992.95	176137
11-01-22	6010	6042.9	5925	5938.75	5938.75	80493
12-01-22	5978.05	6167	5950	6114.1	6114.1	208195
13-01-22	6141	6450	6137	6401.75	6401.75	467887
14-01-22	6411	6415	6232	6312	6312	172383
17-01-22	6350	6505.15	6345	6398.1	6398.1	179777
18-01-22	6480	6523.8	6301	6376.3	6376.3	274448
19-01-22	6547.95	7170	6381.4	6977.25	6977.25	2831099
20-01-22	7059	7525	7037.3	7430.45	7430.45	2156628
21-01-22	7520	7800	7025	7200.7	7200.7	2003434
24-01-22	7275	7495	6834.4	7040.7	7040.7	995137
25-01-22	7085	7280	6840	7179.2	7179.2	805943
27-01-22	7103	7129	6930	6974.15	6974.15	358322
28-01-22	7030.9	7400	7008.05	7183.15	7183.15	603449
31-01-22	7438.7	7777	7376.7	7607.35	7607.35	1006940

The Experimental data for Mind Tree are:

Table 2 Experimental data for Mind Tree

Date	Open	High	Low	Close	Adj Close	Volume
24-Jan-22	4118.85	4118.85	3838.65	3888.95	4118.85	1043403
25-Jan-22	3710	3924.65	3680	3827.2	3888.95	1272983
27-Jan-22	3790	3790	3555	3585.2	3827.2	1469949
28-Jan-22	3646.4	3858.45	3636.55	3724.2	3585.2	1446671
31-Jan-22	3843.95	4070	3835	4006.3	3724.2	1688744
01-Feb-22	4113.9	4117.8	3926.35	3983.7	4006.3	1193072
02-Feb-22	4019.65	4079.35	3973.95	4066.1	3983.7	602678
03-Feb-22	4027	4044	3895.4	3913.1	4066.1	608593
04-Feb-22	3903.1	3988	3828.45	3965.85	3913.1	606538
07-Feb-22	3965	4025	3850.85	3870.05	3965.85	452237
08-Feb-22	3910	3956.4	3821.2	3884.5	3870.05	528879
09-Feb-22	3930	4021.95	3922.05	3993.3	3884.5	605069
10-Feb-22	4024	4041.65	3965.05	4029.55	3993.3	377038
11-Feb-22	3930	3980	3831	3869.65	4029.55	860200
14-Feb-22	3701.5	3835.95	3655.8	3776.45	3869.65	662936
15-Feb-22	3830	3980	3782.9	3967.95	3776.45	718004
16-Feb-22	4030	4036.95	3861.65	3892.2	3967.95	535405
17-Feb-22	3935.5	3935.5	3841.25	3859.45	3892.2	381563
18-Feb-22	3811.05	3879.7	3786	3797	3859.45	378265
21-Feb-22	3796	3871.75	3715	3847.95	3797	562328

It is very important to note that stock price information in the stock market is in volatility, so we should consider characteristics of data while doing the experiment to avoid more learning time, too less training speed, and minimize the error of the prediction result. In this paper, we have concentrated on the unified preprocessing of the data set.

B. Experimental platform construction and parameter settings:

The system configuration mainly include: the CPU type is Intel® core (TM), the hardware device model is HP,64bit Operating System, the memory size is 8GB, the experimental platform is PyCharm 3.1 and the programming language is python language, define the C value as 0.7, gamma as 1, n_epoch as 30

C. Model prediction results:

A stock price prediction model is constructed on the collected stock test data by using SVM method proposed in this paper. Actual value and predictive value of the prediction of the TataElxsi Equity is shown in Figure 4. and MindTree Equity is shown in Figure5.

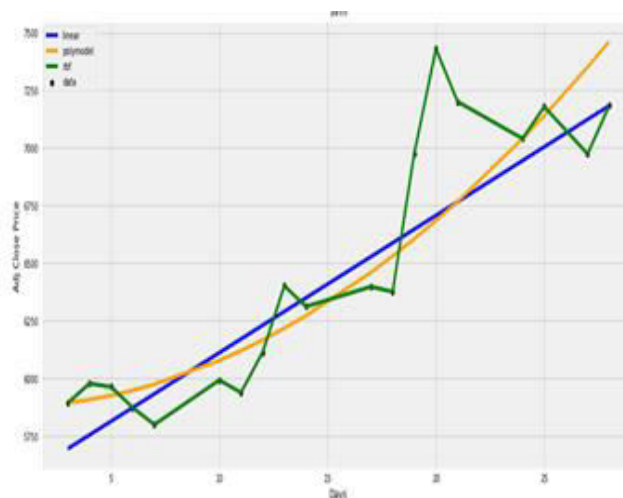


Fig 5.a SVM TATAELXSI

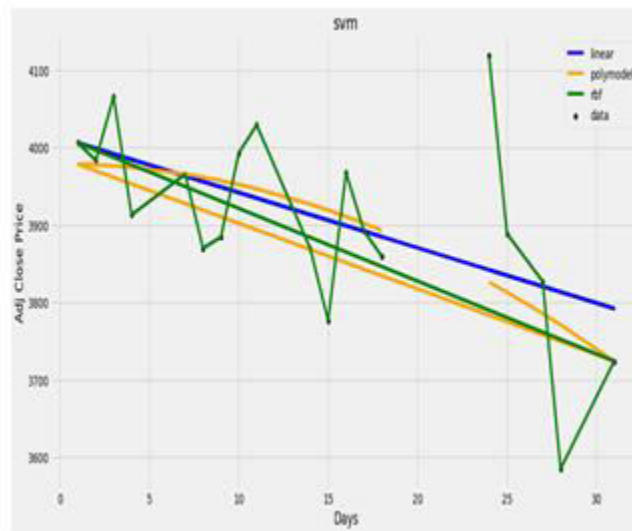


Fig 5.b SVM Mind Tree

More over to test the accuracy of the predicted model, we have examine a comprehensive analysis of the accuracy, standard deviation and running time and their results are in Table 2.

Table 3. Performance of Algorithms for stocks

Method	Error for TataElxsi data set		Error for MindTree data set	
	MSEr	R ² r	MSEr	R ² r
SVM-LINEAR	2.3358	0.8912	2.4281	0.8843
SVM-POLY	2.0358	0.9212	2.0281	0.9143
SVM-RBF	1.3358	0.9912	1.2280	0.9946

Table 3. shows the performance measure of all tested algorithms. In their MSE Error rate of SVM-Linear is higher than SVM-Polynomial and SVM-RBF. R² Error rate is also higher than SVM-POLY and SVM-RBF. Overall, SVM-RBF gives a minimal error rate in both performance measures.

VII. CONCLUSION

This paper presents a comparative study of different SVM kernel model with reliable prediction for closing price of day-to-day stock market. Forecasting stock price is difficult where the situation gives trust for their past results are volatile, whereas this paper gives new ideas with modern technology of machine learning algorithm broke all past results with their optimized accuracy and reliable for closing stock prices. We have to take a past history of stock data to consider as a tool to forecast future close prices. In the research, we have considered stock data of two different companies and implemented with different SVM kernel models. To evaluate these models MSE and R² are used as the performance measures. From the above discussion, we concluded that SVM-RBF has shown higher accuracy than the SVM-Linear and SVM_ Polynomial model.

REFERENCES

[1] M. M. R. Majumder, M. I. Hossain, and M. K. Hasan, "Indices prediction of Bangladeshi stock by using time series forecasting and performance analysis," 2nd Int. Conf. Electr. Comput. Commun. Eng. ECCE 2019, pp. 1-5, 2019.

- [2] L. Sayavong, Z. Wu, and S. Chalita, "Research on stock price prediction method based on convolutional neural network," Proc. - 2019 Int. Conf. Virtual Real. Intell. Syst. ICVRIS 2019, no. DI, pp. 173–176, 2019, doi: 10.1109/ICVRIS.2019.00050.
- [3] Z. Liu, Z. Dang, and J. Yu, "Stock price prediction model based on RBF-SVM algorithm," Proc. - 2020 Int. Conf. Comput. Eng. Intell. Control. ICCEIC 2020, pp. 124–127, 2020, doi: 10.1109/ICCEIC51584.2020.00032.
- [4] F. Liu, X. Li, and L. Wang, "Exploring Cluster Stocks based on deep learning for Stock Prediction," Proc. - 2019 12th Int. Symp. Comput. Intell. Des. Isc. 2019, pp. 107–110, 2019.
- [5] R. Jindal, N. Bansal, N. Chawla, and S. Singhal, "Improving traditional stock market prediction algorithms using Covid-19 analysis," 2021 Int. Conf. Emerg. Smart Comput. Informatics, ESCI 2021, pp. 374–379, 2021, doi: 10.1109/ESCI50559.2021.9396887.
- [6] H. N. Shah, "Prediction of Stock Market Using Artificial Intelligence," 2019 IEEE 5th Int. Conf. Converg. Technol. I2CT 2019, pp. 1–6, 2019, doi: 10.1109/I2CT45611.2019.9033776.
- [7] S. Vazirani, A. Sharma, and P. Sharma, "Analysis of various machine learning algorithm and hybrid model for stock market prediction using python," Proc. Int. Conf. Smart Technol. Comput. Electr. Electron. ICSTCEE 2020, pp. 203–207, 2020.
- [8] G. Bathla, "Stock price prediction using LSTM and SVR," PDGC 2020 - 2020 6th Int. Conf. Parallel, Distrib. Grid Comput., pp. 211–214, 2020, doi: 10.1109/PDGC50313.2020.9315800.
- [9] K. M. Hindrayani, T. M. Fahrudin, R. Prismahardi Aji, and E. M. Safitri, "Indonesian Stock Price Prediction including Covid19 Era Using Decision Tree Regression," 2020 3rd Int. Semin. Res. Inf. Technol. Intell. Syst. ISRITI 2020, no. March 2020, pp. 344–347, 2020, doi: 10.1109/ISRITI51436.2020.9315484.
- [10] X. Kan, M. Miao, L. Cao, T. Xu, Y. Li, and J. Jiang, "Stock Price Prediction Based on Artificial Neural Network," Proc. - 2020 2nd Int. Conf. Mach. Learn. Big Data Bus. Intell. MLBDBI 2020, pp. 182–185, 2020, doi: 10.1109/MLBDBI51377.2020.00040.
- [11] B. B. P. Maurya, A. Ray, A. Upadhyay, B. Gour, and A. U. Khan, "Recursive Stock Price Prediction with Machine Learning and Web Scrapping for Specified Time Period," IFIP Int. Conf. Wirel. Opt. Commun. Networks, WOCN, vol. 2019-December, pp. 1–3, 2019, doi: 10.1109/WOCN45266.2019.8995080.
- [12] Arunachalam, A. S., and T. Velmurugan. "A Survey on Educational Data Mining Techniques." *International Journal of Data Mining Techniques and Applications* 5.02 (2016): 167-171.
- [13] D. Shalinigayathri and A.S. Arunachalam. "Survey on combined supervised learning for optimized daily Stock price". *European Journal of Molecular & Clinical Medicine*, 7, 3, 2020, 5014-5021.