

Impact of using Backend Genetic Algorithm to Optimize Parameters with the use of Support Vector

P. Saiparvathi, V. V. K. Swetha, S. Sandeep, Umamaheswara. P
Department of Computer Science and Engineering
Aditya Institute of Technology and Management
Tekkali, Andhra Pradesh, India

Abstract— This work proposes a new approach, based on Genetic Algorithms and Support Vector Machine to trade in the forex market. In this work, a new algorithm capable of generating technical rules to make investments with a given amount of leverage depending on the certainty of the prediction is presented. To forecast those predictions, a combination of a Support Vector Machine (SVM) algorithm – to identify and classify the market in three different stages –, and a Dynamic Genetic Algorithm – to optimize trading rules in each type of market, is used. The optimization of the trading rules is based on several technical indicators. Forex data for the EUR/USD currency pair, in a timeframe between the years of 2003 and 2016, is used as training and test data. The proposed architecture for the machine learning system, as well as the implementation and study of the proposed system is described in detail. The use of an hybrid system, combining a SVM and a GA with dynamic approaches such as hyper-mutation and adaptability approaches by training three different GA's for each type of market, provide a new approach for FOREX trading, where it is possible to classify trends using price sequences and therefore using the same classification for optimizing investment strategies with the most appropriate GA. Finally, the work shows promising results during the test period between the 2nd of January of 2015 until the 2nd of March of 2016, where the return on investment obtained is 83%.

Keywords— *Forex market; Technical indicators; Support Vector Machine; hyper-mutation; adaptability approaches.*

I INTRODUCTION

A. Genetic Algorithm

In the past, evolutionary algorithms have been applied in many real life problems. GA is one such evolutionary algorithm. GA has emerged as a practical, robust optimization technique and search method. A GA is a search algorithm that is inspired by the way nature evolves species using natural selection of the fittest individuals. The possible solutions to problem being solved are represented by a population of chromosomes. A chromosome is a string of binary digits and each digit that makes up a chromosome is called a gene. This initial population can be totally random or can be created manually using processes such as greedy algorithm [1]

The pseudo code of a basic algorithm for GA is as follows [2]:-

```
Initialize (population)
Evaluate (population)
While (stopping condition not satisfied)
{
    Selection (population)
    Crossover (population)
    Mutate (population)
    Evaluate (population)
}
```

GA uses three operators on its population which are described below:-

Σ Selection: A selection scheme is applied to determine how individuals are chosen for mating based on their fitness. Fitness can be defined as a capability of an individual to survive and reproduce in an environment. Selection generates the new population from the old one, thus starting a new generation. Each chromosome is evaluated in present generation to determine its fitness value. This fitness value is used to select the better chromosomes from the population for the next generation.

Σ Crossover or Recombination: After selection, the crossover operation is applied to the selected chromosomes. It involves swapping of genes or sequence of bits in the string between two individuals. This process is repeated with different parent individuals until the next generation has enough individuals. After crossover, the mutation operator is applied to a randomly selected subset of the population.

Σ Mutation: Mutation alters chromosomes in small ways to introduce new good traits. It is applied to bring diversity in the population.

B. Introduction to Machine Learning

Machine learning is a subfield of artificial intelligence (AI). The goal of machine learning generally is to understand the structure of data and fit that data into models that can be understood and utilized by people.

Although machine learning is a field within computer science, it differs from traditional computational approaches. In traditional computing, algorithms are sets of explicitly

programmed instructions used by computers to calculate or problem solve. Machine learning algorithms instead allow for computers to train on data inputs and use statistical analysis in order to output values that fall within a specific range. Because of this, machine learning facilitates computers in building models from sample data in order to automate decision-making processes based on data inputs[4].

Machine learning is a continuously developing field. Because of this, there are some considerations to keep in mind as you work with machine learning methodologies, or analyze the impact of machine learning processes.

C.Support Vector Machine

SVM stands for Support Vector Machine. “Support Vector Machine” (SVM) is a supervised machine learning which can be used for both classification and regression challenges [10]. However, it is mostly used in classification problems. In the SVM algorithm, we plot each data item as a point in n-dimensional space (where n is number of features you have) with the value of each feature being the value of a particular coordinate. Then, we perform classification by finding the hyper-plane that differentiates the two classes very well

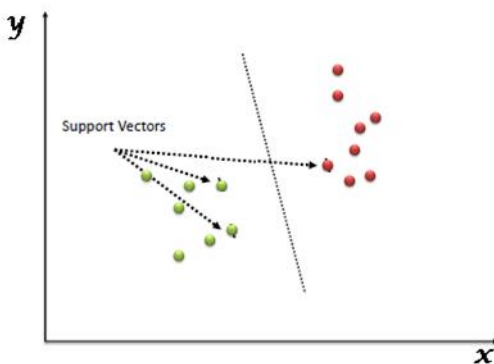


Figure 1-support vector machine representation

Support Vectors are simply the co-ordinates of individual observation. The SVM classifier is a frontier which best segregates the two classes

II RELATED WORK

A. Introduction to Forex Market

Forex Market is an exciting place. The one good thing about entering into the forex market is that you can trade anytime as per your convenience.

The global Foreign exchange market (‘FX’, ‘Forex’ or ‘FOREX’) is the largest market in the world as measured by the daily turnover with more than US\$5 trillion a day eclipsing the combined turnover of the world’s stock and bond markets. The forex market measuring a propelling turnover is one of the many reasons why so many private investors and individual traders have entered the market. The investors have discovered several

Forex (in simple terms, currency) is also called the foreign exchange, FX or currency trading. It is a decentralized global market where all the world’s currencies trade with each other. It is the largest liquid market in the world. The liquidity (more buyers and sellers) and competitive pricing (the spread is very small between bid and ask price) available in this

marked are great. With the irregularity in the performance in other markets, the growth of forex trading, investing and management is in upward trajectory[5].

So, why trade Forex? There are many reasons to trade in Forex. If we ask four different people, you might get more than four different answers. Primarily, making money is the most frequently cited reason for why trade Forex. Let us now consider the following reasons why so many people are choosing forex market –

Forex market never sleeps-- The Forex market works 24 hours and 5-1/2 days a week. Because governments, corporates and private individual who require currency exchange services are spread around the world, so trading on the forex market never stops. Activity on the forex market follows the sun around the world[6], so right from the Monday morning opening in Australia to the afternoon close in New York. At any point of the day you can find an active pair to trade.

Long or short-- A trader in forex can trade both ways. It means a forex trader can play the market and make profits irrespective of whether market is going up, down or is in tight range. So irrespective of the event that has triggered the movement – forex traders do not care.

Low transaction cost-- Most forex accounts trade with little or no commission and there is no exchange or data license fees. Generally, the retail transaction fee (the bid/ask spread) is typically less than 0.1% under normal market conditions. With larger dealers (where volumes are huge), the spread could be as low as 0.05%. Leverage plays a crucial role here.

Leverage-- Leverage is the mechanism by which a trader can take position much larger than the initial investment. Leverage is one more reason why you should trade in forex. Few currency traders realize the advantage of financial leverage available to them. For example, if you are trading in equity market, the maximum leverage a stock broker is offered is 1:2 but in case of forex market, you will get a leverage up to 1:50 and in many parts of the world even higher leverage is available. For this reason, it is not hard to see that why forex trading is so popular.

High leverage allows a trader with small investment to trade higher volumes of currencies and thus provide the opportunity to make significant profits from the small movement in the market. However, if the market is against your assumption you might lose significant amount too. Therefore, like any other market, it is a two-way sword.[7].

B.Algorithm

Genetic Algorithm:

(a)Selection:

For individuals in population:

Grouping the individuals in tourments of size

k

for group made:

choose individuals with the best fitness within the group

select best individuals from each group

(b)Mutation:

If random probability < mutation probability:

For gene in chromosome:

If random probability < gene probability:

mut_gene=mutate gene with Gaussian function
 addmut_gene and gene to return the mutate gene
 Support Vector Regression:

Required:

X and Y are loaded as a training, c as a regulation factor and epsilon>0

1: clf<= SVR(kernel='rbf', degree=3, epsilon = E)

2: repeat:

for all {xi,yi},{xj,yj} do

Optimize α and α_j

End for

Until no change in α

III.EXPERIMENTAL RESULTS

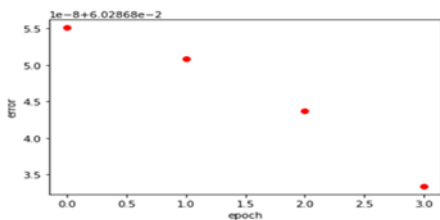
A. Dataset Description

Forex dataset contains 7 columns which indicates the notations of currency belonging to different countries. These datasets having currency values up to 10 years of predicted values which helps in predicting the present values. The notations used in the dataset are SP, DAX, FTSE, NIKKEI, BOVESPA, EU, and EM.

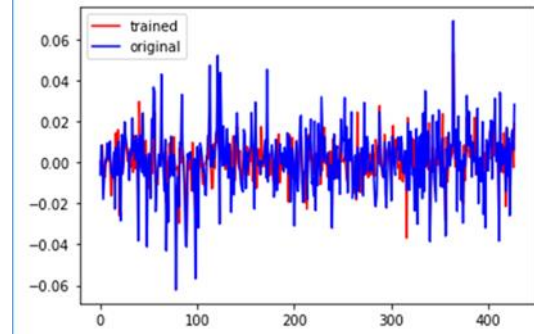
- NSE: National Stock Exchange
- SP: S & P 500 INDEX (New York Stock Exchange)
- DAX: DeutscherAktien Index (Frankfurt Stock Exchange)
- FTSE: FTSE 100 Index (London Stock Exchange)
- NIKKEI: Nikkei Index (Tokyo Stock Exchange)
- BOVESPA: Bovespa Index (Brazil Sao Paulo Stock Exchange)
- EU: MSCI Europe Index
- EM: MSCI Emerging Markets Index
- MSCI: Morgan Stanley Capital International.[11]

B. Outputs

```
current error: 0.06028683861211882
current error: 0.06028683780082175
current error: 0.06028683689767736
train var : 0.4260887503549663
train r2 : 0.4260887502708871
train rmse : 0.011868322408295223
0:00:03.422036
test var : 0.42599528944952214
test r2 : 0.42373463336038564
test rmse : 0.013812010172431066
```

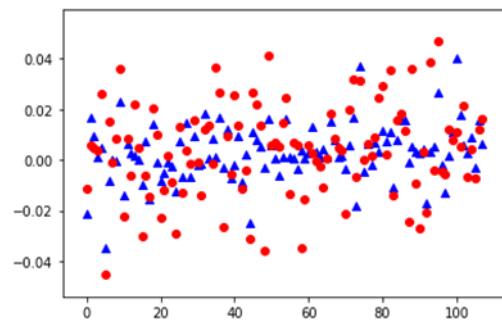


In the above figure, the graph shows the error rate for analysis of stock market index with the Gaussian functions. The graph is plotted between the current error and the complete representation of dataset (epoch).



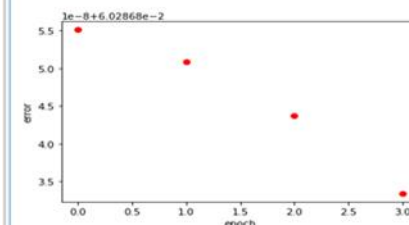
I.

This figure represents the noise formed during comparison of training data and the original data. The training data is the outcome of support vector machine in which it gets data for genetic algorithm by splitting the original dataset by selection and mutation.

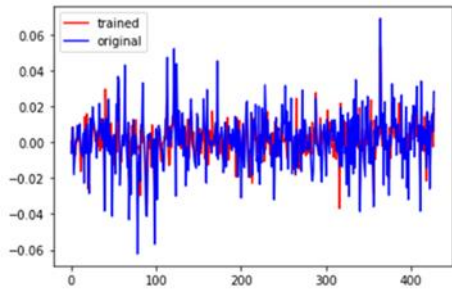


This figure is the just analysis of same result with ANFIS algorithm. This is for observing that the predicted values and analysis of the stock indexes.

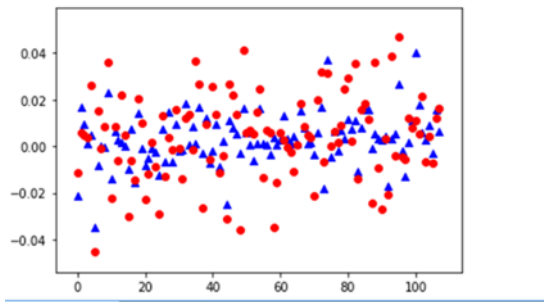
```
current error: 0.06028683861211882
current error: 0.06028683780082175
current error: 0.06028683689767736
train var : 0.4260887503549663
train r2 : 0.4260887502708871
train rmse : 0.011868322408295223
0:00:03.422036
test var : 0.42599528944952214
test r2 : 0.42373463336038564
test rmse : 0.013812010172431066
```



This figure is the 16 value of the complete output, similar to the value 1 as we calculate error rate for some iterations until the error rate becomes negligible.



Similar to the previous output the comparison will be done for every iteration and plotted into noisy graph. Blue for training data and red for original data which are taken from support vector machine outcome.



In the above figure is the ANFIS representation of analysis of the output for the value 16. The difference between training data and test data.

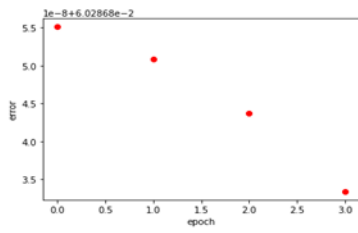
```

0      ISE.TL
1      0.035753708
2      0.025425873
3      -0.02886173
4      -0.062208079
5      0.009859905
    
```

In the above figure there are some predicted tl (turkey index value), which we want the output. the input values are forex dataset after analysis of the dataset indexes we predict value of one index value.

```

current error: 0.060286818288988235
current error: 0.060286619844657434
current error: 0.060286582295631986
train var : 0.4260949108068872
train r2 : 0.4260945692485132
train rmse : 0.011868262240736525
0:00:03.434634
test var : 0.426006584809528
test r2 : 0.4236989553790127
test rmse : 0.013812437733276332
    
```



Finally this figure is the last iteration value of current error, this also having the same process and representation as the

previous values when the error is same for some iterations then it is finalized as the best solution.

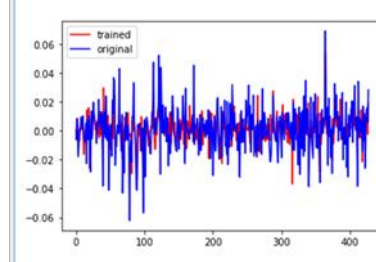
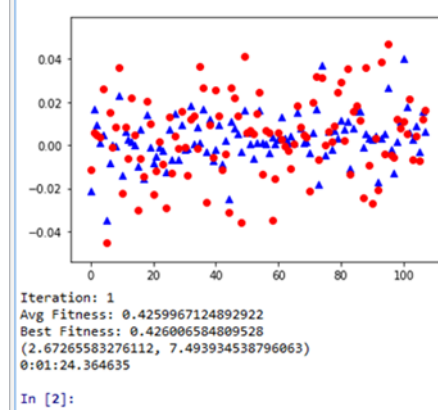


Figure show the final value of support vector machine by comparing the training data and original data in a noisy data graph representation.



This figure shows the analysis of the all outcomes in ANFIS algorithm. This gives the average fitness as the best solution for predicting the index values.

IV CONCLUSION

This work proposes a viable solution to automatically invest in Forex Markets, using technical indicators and price sequences to predict entry and exit points, although, it is highly recommended to use hardware that has good performance measures, so that is possible for the algorithm to work in real time . There commendation is made because, the more information about the FX market, the bigger the populations and the more price sequences the SVM is provided to improve the training module, the better the classification of the market is .Also, by studying the SVM approach, it is possible to conclude by the results, that the sideways market is more difficult to describe than the others as it was seen in chapter 4, where the label “0” showed poorer quality in the classification, due to the fact that the Sideways market is difficult to be classified even by a human eye. Both Bullish and Bearish markets were classified with very successful scores of precision, recall and accuracy. By the results obtained in the test period it is possible to conclude that the classification a market is highly important, as an environment as well as a tool that allows the implementation of leverage. The environment classification passed from the SVM on to the three different GAs obtained very promising results, since the Proposed System presented a steady growth with a low Drawdown. In a more technical matter, the SVM proved to be a reliable classifier, said that it is important to address the following conclusions:

- The way data is organized, is the key to get a good SVM classifier, i.e., for most problems, and especially the one addressed in this thesis, combined a great amount of data that is not suitable to use directly from the index. An organized and well-structured dataset improves the performance and the quality of the results that are obtain by the SVM model.
- The quality of the training set is also a very important issue, since this data set is the foundation on which the SVM learns. If the data set is not well balanced between the different labels, the metrics may return ill-defined results that do not translate the reality of the classification. Another measure for data set quality is whether the training set is correlated with the testing set, e.g., if the training set only has data from a FX chart that dates from 1999, and the test set is from the year 2015, then probably, the classification will be ill-defined since the data is not highly correlated anymore, since the FX market has changed a lot in 16 years.
- The cross-validation used is a very important tool to assess if the training set is not over fitted before testing the model on the test set. This tool proved to be very useful to develop the model, in order to ensure that the model, did not “leak” information between the training and test set, i.e., if the SVM model has any information about future actions, then the results cannot be reliable.

V REFERENCES

- [1] M. Mitchell, "Genetic Algorithms: An Overview," *Complexity*, Vol. 1, No. 1, pp. 31–39, Sep. 1995.
- [2] A. Neubauer, "Theory of The Simple Genetic Algorithm with \square -Selection," In *Proceedings of The 10th Annual Conference On Genetic and Evolutionary Computation (GECCO '08)*, Atlanta, USA, New York: ACM, 2008, pp. 1009–1016.
- [3] P. K.-F. Man, K. S. TANG, and S. Kwong, *Genetic Algorithms: Concepts and designs, Illustrated*. Springer Science & Business Media, 2012.
- [4] J. Holland, "Adaptation in Artificial and Natural Systems." Ann Arbor: The University of Michigan Press, 1975.
- [5] A. E. Drake and R. E. Marks, "Genetic Algorithms in Economics and Finance: Forecasting Stock Market Prices and Foreign Exchange — A Review," *Genetic Algorithms and Genetic Programming in Computational Finance*, pp. 29–54, Jan. 2002.
- [6] R. W. Colby and T. A. Meyers, *The Encyclopedia of Technical Market Indicators*. New York: Irwin, 1988.
- [7] Investopedia.com, "Leverage Definition," Investopedia, 2003. [Online]. Available: <http://www.investopedia.com/terms/l/leverage.asp?layout=orig>.
- [8] P. H. Winston, "Lecture 13: Learning: Genetic Algorithms," MIT Open Courseware. [Online]. Available: <http://ocw.mit.edu/courses/electrical-engineering-and-computer-science/6-034-artificial-intelligence-fall-2010/lecture-videos/lecture-13-learning-genetic-algorithms/>.
- [9] L. Martí, "Luis Martí," 2014. [Online]. Available: <http://lmarti.com/genetic-algorithms-algoritmos-geneticos-eng1456>.
- [10] S. Kei, D. Shangkun, and S. Akito, "Prediction of Foreign Exchange Market States with Support Vector Machine," In *Machine Learning and Applications and Workshops (ICMLA)*, 2011, 10th International Conference On, IEEE, Vol. 1, pp. 327–332.
- [11] <http://www.kaggle.com/uciml/istanbul-stoch-exchange>.