GENETIC ALGORITHM BASED OPTIMIZATION OF SELECT TECHNICAL TRADING RULES: EVIDENCE FROM INDIAN STOCK MARKET

Jayanthy B, Dr. Tamilchelvi N

Abstract— The objective of this research is to optimize the select technical trading rules. The rules uses a genetic algorithm based on technical analysis indicators such as Exponential Moving Average (EMA), Moving Average Convergence Divergence (MACD), Relative Strength Index (RSI) and Filter that gives buying and selling recommendations to investors. The algorithm optimizes the strategies by dynamically searching to improve profitability in the training period. The best sets of rules are then applied on the testing period. This paper proposes a novel idea to find a suitable technical trading rule which is systematically optimized to generate excess returns in the leading Indian Stock markets.

I. INTRODUCTION
This paper describes a genetic algorithm used to create a trading system, consisting of several rules for opening and closing trading positions in the Indian market. The aim of this study is to assess the profitability of theNSE/BSE market. An investor seeking to make a profit in the stock market, should develop alternate models dedicating a fair amount of resources. In this study the data and resources are from the leading Indian Stock markets BSE andNSE. The findings capture the attention of investors in the Indian Stock market that base their decisions on technical analysis signals.

Chapter organization
The research proceeds as follows. This section offers a review of the existing literature related to the study. Section 2 presents the database used for testing the trading rules and the methodology involved. Section 3 discusses empirical findings of the analysis and Section 4 concludes and offers future directions.

Technical analysis is the methodology of forecasting price movements by analyzing past market data [1] Most automated trading systems use several indicators in order to generate purchase and sale recommendations [2]. In [3] it is found that one can achieve performing returns using trading strategies only if he has full information of the stock price change for the future. However, if the future information is not accurate, it can be useless in increasing profits. Moreover, a search in a strategy space to get high profit is impossible and this is based on lack of future information of a company. Trading strategies have been mainly based on technical analysis in the commodity futures market [4], [5], [6] and foreign exchange market [7], [8], [9], [10]. Evaluation of the technical analysis’ performance in the equity markets has generally been done using market indices. Technical analysis has evolved beyond filter and moving averages rules, now including psychological barriers such resistance and support levels [11]; [12].

Genetic algorithms (GA) represent a class of optimization techniques that generate solutions to search problems and quickly adapt to changing environments. GA were developed by Holland [13] and they simulate the process of natural evolution. As the species evolve through genetic processes such as selection, crossover and mutation, GA create classes of solutions that evolve over some generations through analogous processes in order to generate one solution with the best fit to the specific problem [14]. Algorithms start by creating some strategies with specific parameters. In the following steps, they dynamically change their parameters in order to achieve higher profits. Holland [13] developed a way in which the natural evolution process might be imported into algorithms that offer solutions to search problems. GA are very suitable for managing financial markets because these represent a continuous changing environment and trading strategies need to adapt to the new conditions. Another study about the performance of the genetic algorithms for Indian markets has been developed in [15]. The authors show that the system often returned profit when the testing period was consecutive to the training period. They concluded that the success of the system depended on the similarity in the trends of the two periods. Also, genetic algorithms succeeded in finding performing trading rules for six exchange rates [16].

II. Research Methodology
The database used in this work is the tick-by-tick series ofNSE/BSE currency pair over the year 2013. Time series with frequency of 60 minutes were used for testing the performance of the genetic algorithm. Time series have been separated in two sets: the training period and the testing period. The first one considers the first six months of 2013 and is used for finding the strategy that achieves the highest performance. The second set tests the performance of the strategy found in the first step. The algorithm is applied 150 times on the training time series, in order to find the characteristics of the best 150 individuals. Then the performance of these individuals on the out-of-sample series is assessed. The hourly data extracted from the tick-by-tick data also consider the minimum and maximum tick for both bid and ask quotes. This information is needed to establish if the take-profit or stop-loss level had been reached during that period.

Manuscript received Sep 01, 2014
Jayanthy B, Assistant Professor, Department of Commerce, Nirmala College for Women, Coimbatore, India
Dr. Tamilchelvi N, Associate Professor and Head, Department of Commerce, TSAAS. Tamil College, Perur, Coimbatore, India

33 www.ijerm.com
GENETIC ALGORITHM BASED OPTIMIZATION OF SELECT TECHNICAL TRADING RULES: EVIDENCE FROM INDIAN STOCK MARKET

The purpose of the genetic algorithm is to optimize a set of trading rules to generate higher profit. Trading rules base their decisions on several indicators presented below together with their formulas.

**Exponential moving average.** It gives greater weights to the latest prices when computing the average. When the price is above this indicator, the signal is for a long (buying) position and when the price is below the exponential moving average it signals the selling.

\[
EMA(n, \text{close}) = \left\{ \frac{2}{n+1} \right\} \cdot \text{Close}_t + \text{EMA}_{t-1} \cdot \left\{ 1-\frac{2}{n+1} \right\}
\]

Where \( EMA \) is the exponential moving average indicator, \( n \) is the number of periods and \( \text{Close} \) is the closing price of the period.

**Moving Average Convergence Divergence** – it is an indicator based on several other technical analysis indicators. Firstly, the \( MACDline \) is computed as a difference between two exponential moving averages. Secondly, we compute a signal line as an exponential moving average of the \( MACDline \). Finally, the MACD indicator is computed as the difference between the \( MACDline \) and the signal.

\[
MACDline(p,q) = EMA(p,\text{close}) - EMA(q, \text{close})
\]

\[
\text{Signal}(m) = EMA(m,MACDline(p,q))
\]

\[
MACD(p,q,m) = MACDline(p,q) - \text{Signal}(m)
\]

Where, \( p \) is the number of periods of the short exponential moving average, \( q \) is the number of periods of the long exponential moving average, \( \text{Close} \) is the closing price of the period and \( m \) is the number of periods of the moving average of the \( MACDline \).

This indicator offers buying or selling signals when its value is positive, respectively negative.

**Relative Strength Index** – is a technical analysis indicator that gives overbought and oversold signals. The overbought signal is given when the RSI value is over a specific benchmark (usually 70 or 80) and the oversold signal is given when this value is under another benchmark (the standard is 20 or 30).

\[
RSI(p) = \frac{1}{n} \sum_{i=1}^{n} \frac{\max(P_i-P_{i-1},0)}{P_{i-1}}
\]

Where \( P \) is the closing price of the period and \( n \) is the number of the periods used to compute the RSI.

Further, we start the description of the algorithm with the definition of an individual.

**The Individual Characteristics**

In a genetic algorithm for setting a FX trading system, each individual is represented by a set of technical analysis rules. Each rule can be considered as a chromosome, while the parameters that define a rule are considered genes. Here we consider the individual as being defined by 6 chromosomes (rules) and 24 genes (parameters). The rules are divided in 4 rules that set the conditions for opening a position and the rest 2 rules are those that define the conditions for exiting the position. Each rule contains a Boolean gene that can activate or deactivate the rest of the rule’s genes.

Following, are described the rules (chromosomes).

**Rules for Position Opening**

**Rule 1. Exponential Moving Average: EMA(n).** Genes:

- **Boolean** \( EMA \) takes the values 0 or 1. Value 0 deactivates the rule, while value 1 activates it.
- **Nr_periods_EMA** , noted \( n \), takes values between 5 and 90.

The rule generates trades as follows. If there is no current open position, then a long position is generated only if the close price is higher than EMA(n) and a short position is generated only if the close price is lower than EMA(n). If a position is currently open, then this rule is ignored.

**Rule 2. Moving Average Convergence Divergence:**

**MACD(p,q,m).** Genes:

- **Boolean** \( MACD \) – takes the values 0 or 1. Value 0 deactivates the rule, while value 1 activates it.
- **Periods_short_MA**, noted \( p \), takes values between 5 and 90
- **Periods_long_MA**, noted \( q \), takes values between 10 and 100, with the restriction \( q > p \)
- **Periods_signal_MACD**, noted \( m \) - takes values between 5 and 25 and is the moving average of the difference between the short and the long moving average.
- **Boolean** \( signal \) - takes the values 0 or 1. Value 0 sets the value for the signal to 0. Basically, it transforms the MACD in a simple rule of moving averages crossover. Value 1 activates the signal. For the 0 value is attached a probability of occurrence of 25%, while for value 1 the probability is set to 75%

PerioIf there is no currently open position, then, if the **Boolean_signal** has the null value, the rule takes into en the short moving average is higher than the long one and a short position is opened otherwise. Consideration only the short and the long moving averages. Therefore, a long position is opened wHIf the **Boolean_signal** takes the value 1, the rule proceeds as follows. If the difference between the short moving average and the long one is higher than the value of the signal, then a long position is opened, while otherwise is opened a short position.

**Rule 3. Relative Strength Index: RSI (n).** Genes:

- **Boolean** \( RSI \) – takes the values 0 or 1. Value 0 deactivates the rule, while value 1 activates it.
- **Periods_RSI** , noted \( n \), takes values between 5 and 50.
- **Oversold_signal_RSI**, noted \( p \), takes values between 15 and 35.
- **Overbought_signal_RSI**, noted \( q \), takes values between 65 and 85.
- **Boolean_signal_RSI** - takes the values 0 or 1. The use of this gene is described below
The rule generates trades only if currently there is no open position. The trades are generated based on the Boolean_signal_RSI value as follows: When it takes the value 0, a long position is opened when the RSI value drops under \( p \) and a short position is opened when the RSI value rises over \( q \). When it takes the value 1, a long position is opened when the RSI value rises over \( p \) and a short position is opened when the RSI value drops under \( q \).

**Rule 4. Filter(n). Genes:**

- Boolean_Filter – takes the values 0 or 1. Value 0 deactivates the rule, while value 1 activates it.
- Periods_Filter, noted \( n \), takes values between 1 and 15.
- Increase_signal_Filter, noted \( p \), takes values between 50 and 100 pips.
- Decrease_signal_Filter, noted \( q \), takes values between 50 and 100 pips.
- Boolean_trader_type_Filter - takes the values 0 or 1.

This rule respects the same restriction as the rest of the three opening rules: if there is already a currently open position, the rule is ignored. The trades are generated based on the Boolean_trader_type_Filter value as follows: Value 0 signals a trend follower (enters long if the price increases more than \( p \) pips or short if the price decreases more than \( q \) pips). Value 1 signals that the trader will enter long if the price decreases more than \( q \) pips or short if the price increases more than \( p \) pips.

For the above, a great importance have the Boolean genes that activate or deactivate the rules: 1, 3, 8 and 13. When all of them take the value 0, this means that the individual will never open a position (because no opening rule is active). In order to avoid such situations, that have a probability of occurrence of 6.25%, we proceed the following way. If these genes take all value 0, then we randomly change the value for one of them. Moreover, if two or more of these genes take simultaneously the value 1, then a position is opened only if all the active rules give the same trading signal (to buy or to sell). Therefore, it is more probable that an individual that has only one active rule to trade more than an individual that has all the rules active.

As important as the rules that define the conditions to open a position are the rules used to exit that position, in order to take the profit or to cut the losses. Following are described these rules.

**Rules for Exiting the Position**

**Rule 5. Fixed exit levels \((p,s)\). Genes:**

- Boolean_fixed_exit – takes the values 0 or 1. Value 0 deactivates the rule, while value 1 activates it.
- TP_fixed, noted \( p \), takes values between 15 and 150 pips
- SL_fixed, noted \( s \), takes values between 10 and 100 pips

Opposite to the opening rules, the rules for exiting the position are active only when a position is opened. The above rule exits a long position if the price rises at least \( p \) pips (take profit) or drops at least \( s \) pips (stop loss). Accordingly, the rule exits a short position if the price drops at least \( p \) pips (take profit) or rises \( s \) pips (stop loss).

**Rule 6. Trailing exit levels \((p,s,q)\). Genes:**

- Boolean_trailing_exit - takes the values 0 or 1. Value 0 deactivates the rule, while value 1 activates it. This gene is conditioned by gene number 18. If gene 18 takes value 0, then gene 21 takes value 1 and if gene 18 takes value 1, then gene 21 takes value 0.
- TP_trailing, noted \( p \), takes values between 15 and 150 pips
- SL_trailing, noted \( s \), takes values between 10 and 100 pips
- Trailing_level, noted \( q \), takes values between 10 and 100 pips, under the restriction that \( q < p \).

The above rule can be active only if a position is already open and rule 5 is not active. If a long position is already open and the price rises at least \( q \) pips, but less than \( p \) pips, the take profit and stop loss levels are updated, by increasing them with \( q \) pips. Continuing, if the price rises another \( q \) pips, but the new take profit level is not reached, then the stop loss and take profit levels are updated again, by increasing them with another \( q \) pips. The same procedure is followed until the stop loss is reached or during one period the take profit is hit. In the case of a short position, same methodology is used, with the difference that the stop loss and take profit levels are updated by decreasing them with \( q \) pips.

**III. The Genetic Algorithm**

After defining the individual, characterized by the rules for entering into position (based on the technical analysis indicators) and by the exit rules (take profit and stop loss), we proceed to the genetic algorithm, which consists in the following steps:

1. A population of 100 individuals is randomly generated.
2. We compute the profit or loss generated by each individual over the training period. Each individual is evaluated based on this measure.
3. The individuals are ranked based on the generated profit or loss.
4. The new generation is created by the following procedures:
   - **A. Elitism.** The most profitable individual is automatically passed to the new generation
   - **B. Selection of parents.** The probability of a given individual to become a parent for the new generation is based on its ranking. In order to increase the computational speed, we divided the individuals in 10 classes of fitness (profitability). First class contains the first 10 best-ranked individuals, the second class contains the individuals ranked 11th to 20th, while the 10th class contains the last 10 ranked individuals (Table 1). For the individuals of the same class, we attach the same probability. In addition, the probability is higher for classes that contain better-ranked individuals (e.g. the first class will
have attached a higher probability than the 10th class.

C. Crossover. Using the selection criteria described above, pairs of two parents are randomly chosen. Each pair of parents will create a new individual. In order to choose what genes from what parent will be passed to the new individual, a number \( n \) (where \( 1 < n < 24 \)) is randomly generated. The new individual will receive the genes 1 to \( n \) from one parent and the genes \( n+1 \) to 24 from the other parent. The gene 21 will still depend on the gene 18. This way 80 individuals from the new generation are obtained.

D. Introduction of migrants. In order to increase the diversity and to avoid fast convergence, we randomly generate 19 individuals in the new generation.

| Table 1 | Probabilities of individuals to become parents. |
|---|---|---|---|---|
| class | No of individuals in the class | Probability for each individual in the class | Class Probability | Cumulative Probability | Cumulative Population |
| 1 | 10 | 2.5% | 25% | 25% | 10 |
| 2 | 10 | 2.2% | 22% | 47% | 20 |
| 3 | 10 | 1.2% | 12% | 59% | 30 |
| 4 | 10 | 1.00% | 10% | 69% | 40 |
| 5 | 10 | 0.8% | 8.0% | 72% | 50 |
| 6 | 10 | 0.6% | 6.0% | 78% | 60 |
| 7 | 10 | 0.34% | 3.4% | 81.4% | 70 |
| 8 | 10 | 0.3% | 3.0% | 84.4% | 80 |
| 9 | 10 | 0.25% | 2.5% | 86.9% | 90 |
| 10 | 100 | 0.155 | 15.5% | 100.0% | 100 |

Notes: Table 1 shows the probabilities of individuals to become parents in the crossover step of the genetic algorithm. The probabilities are sorted by the cumulative profitability in the training period and, in order to increase the computational speed, the individuals are divided into 10 classes of fitness.

5) The new generation becomes the actual generation and the steps 2–4 are repeated.

6) We repeat the procedure from steps 2–5 until we reach 30 such iterations.

By executing the genetic algorithm, one individual, the result of the evolution. We repeat the genetic algorithm for 100 times in order to obtain 100 such individuals (sets of trading rules). Then, these 100 best individuals from the training period will be evaluated on the testing period. The evaluation procedure consists in assessing the profit or loss (expressed in pips) generated by the each individual in the testing period. The results obtained are attached in the Supporting Information file.

Results and Discussion
To analyze the results, we firstly discuss the evolution of NSE/BSE in the training and testing period (Fig. 1). During the training period, a short upward movement, followed by a sideways evolution, firstly characterizes the exchange rate. Starting with May 2012, a strong downward trend is set. The testing period starts with a continuation of the downward trend, followed by a reversal and an upward trend in August 2012. The final part of the testing period is characterized by a sideways evolution of the NSE/BSE exchange rate. Both training and testing period contain price movements in trend or sideways. Therefore, it is expected that the rules that perform relatively well in both types of markets (trending and sideways) will obtain good results in both periods.

![Figure 1] NSE/BSE over the period 2012.

The cumulative profit exhibits an upward trend on the training period for all the 100 best individuals (Fig. 2). The increase in the cumulative profit does not have important variations, showing that the individuals are well fitted on the training period. However, on the testing period, the cumulative profit seems uniformly distributed around the null value and its dispersion increases with time (Fig. 3). The individuals that performed best on the training sample are not able to achieve similar results on the testing sample, providing evidence that NSE/BSE market is weak-form efficient.

![Figure 2] Profit on the training period – simulation 1.
Two more simulations are made of the program in order to verify the consistency of this proposed results and the parameters of the generated individuals are attached in the Materials S1 file, together with those of the first simulation. In the case of the second simulation, the results for the training period are very similar to those obtained in the initial one (Fig. 4). In addition, the cumulative profit over the testing period exhibits the same pattern of the first simulation (Fig. 5). By running the third simulation, the results are very similar (Fig. 6, Fig. 7). Therefore, these simulations validate the initial results that the best performers over the training period are not able to achieve similar results over the testing period. Our results are consistent with those obtained by Mendes et al. [34], suggesting the weak-form efficiency of the NSE/BSE market.

Next, the statistics of all the 300 generated individuals for the training and testing periods are computed. Statistics with and without transaction costs are computed. Results are similar in both cases. Statistics on the training sample show that the minimum, maximum and average cumulative profit are all positive and high (Table 2). This happens because each selected individual is the most profitable from a set of 3000 individuals. Therefore, their outcome is predictable high.

### Table 2 Statistics for the cumulative profit on the training period

<table>
<thead>
<tr>
<th>Type of profit</th>
<th>Average</th>
<th>Median</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Std. dev</th>
<th>Nr Profitable</th>
<th>Skewness</th>
<th>Kurtosis</th>
<th>JB</th>
<th>P val</th>
</tr>
</thead>
<tbody>
<tr>
<td>Real</td>
<td>1,900.0</td>
<td>1,876.1</td>
<td>1,314.00</td>
<td>2,898.00</td>
<td>284.15</td>
<td>300</td>
<td>0.49</td>
<td>2.95</td>
<td>12.25</td>
<td>0.002</td>
</tr>
<tr>
<td>Without Cost</td>
<td>2,007.8</td>
<td>1,968.6</td>
<td>1,436.40</td>
<td>2,967.70</td>
<td>275.93</td>
<td>300</td>
<td>0.57</td>
<td>3.13</td>
<td>16.33</td>
<td>0.000</td>
</tr>
</tbody>
</table>

Notes: Table 2 shows the statistics for the cumulative profit expressed in pips on the training period. The statistics are computed for both cases: with and without transaction costs.

The second period is a robustness test for the strategies found in the first period. The average cumulative profit at the end of the testing period is negative, but close to 0, being consistent
with efficiency hypothesis that no arbitrages can be made using the winning strategies from period 1 (Table 3).

In addition, the variability of the outcomes is higher in the testing period (the standard deviation is almost double in the testing period than in the training one). The values of the Skewness and Kurtosis statistics provide evidence that the profit distribution over the testing period may be normal. The empirical distribution plotted in Fig. 8 shows that the profits follow a distribution close to the normal one, but it is skewed from the standard normal distribution due to its negative average.

![Distribution of the profits on the testing period.](image)

**Table 3 Statistics for the cumulative profit on the testing period.**

<table>
<thead>
<tr>
<th>Type of profit</th>
<th>Average</th>
<th>Median</th>
<th>Minimum</th>
<th>Maximum</th>
<th>St. dev</th>
<th>N Profitable</th>
<th>Skewness</th>
<th>Kurtosis</th>
<th>JB</th>
<th>P val</th>
</tr>
</thead>
<tbody>
<tr>
<td>Real</td>
<td>−198.9</td>
<td>−224.0</td>
<td>−1,780.0</td>
<td>1,702.0</td>
<td>550.86</td>
<td>107</td>
<td>0.07</td>
<td>3.04</td>
<td>0.26</td>
<td>0.8788</td>
</tr>
<tr>
<td>Without Cost</td>
<td>−109.9</td>
<td>−130.8</td>
<td>−1,672.9</td>
<td>1,749.60</td>
<td>560.75</td>
<td>126</td>
<td>0.05</td>
<td>2.92</td>
<td>0.22</td>
<td>0.8978</td>
</tr>
</tbody>
</table>

Notes: Table 3 shows the statistics for the cumulative profit expressed in pips on the testing period. The statistics are computed for both cases: with and without transaction costs.

Further, we have applied Jarque-Bera test in Eviews 7 to check for normality. The JB statistic is computed as:

\[
JB = \frac{n}{6}(S^2 + (1/4)(K-3)\bar{K})\] ...(4)

Where, \( n \) is the number of observations, \( S \) is the skewness and \( K \) is the kurtosis. The test is built on a joint null hypothesis of a skewness of 0 and a kurtosis of 3 because these values characterize the normal distribution. As shown in Table 3, the null hypothesis of normal distribution for the testing period profits cannot be rejected.

A frequent problem met in the case of technical trading rules is the data-snooping bias. It may appear when more strategies are tested on the same sample. In this way, a rule may be performing in a period only due to luck. Therefore, when it is applied to another period, it generates negative returns. In the literature, a data-snooping test is applied to check for the validity of good performance. In our case, the out-of-sample results are distributed around 0, showing that in the case of the NSE/BSE market one cannot find an outperforming strategy based on historical prices. Therefore, in the absence of a consistently profitable strategy (genuine or due to luck), the data-snooping test is not needed in this proposed algorithm.

Concluding, this proposed results show that the hypothesis of weak-form efficiency cannot be rejected in the case of NSE/BSE market. Of course, this does not necessarily mean that one cannot prove the market inefficiency by finding a set of rules that consistently achieve profits. However, finding this set of rules represents a difficult task. We consider that our main results suggest that an investor should carefully analyze before taking speculative positions based on technical indicators and computer-based algorithms because there are higher chances to lose on the long-run. The fact that a sophisticated algorithm was not able to achieve sustainable profits supports our remark.

The future research adding some filters to the trading strategies in order to avoid false signals is recommended. For example, a strategy may achieve better results if the investor enters a position only after receiving the same signals for several periods. The same filter can be applied for the exit rules. Moreover, if some strategies are found to be performing, a data-snooping test should be applied in order to check their genuine predictive power.

**Reference**


