MARKET SPECIFIC TECHNICAL INDICATORS COMBINING THE WISDOM OF CROWD

Beenish Bashir  
COMSATS University Islamabad. Corresponding Email: beenish128@hotmail.com  
Faheem Aslam  
COMSATS University Islamabad. fahimparacha@gmail.com  
Aneel Salman  
COMSATS University Islamabad. aneelsalman@gmail.com

ABSTRACT

To improve the profitability and predictability of financial markets we highlight the issue of relevant technical indicators (TI) for different markets. In previous studies, not much attention has been given to feature selection problem by counting on the most popular technical indicators or following the footsteps of related literature. This research work is focusing on the importance of feature selection problem. There is no specific set of TI that is good enough to predict every market price movement. TI which are good predictors of developed markets are not suitable for emerging or frontier markets. The analysis is based on a well-diversified sample of nine countries representing developed, emerging and frontier markets as per the categorization of Morgan Stanley Capital International (MSCI). We use random forest (RF) technique for feature selection from a group of 90 technical indicators. The results show that top five technical indicators are different for each market according to Gini index. Even within the three categories of market, these indicators and their ranking varies. Hence, appropriate feature selection according to market eventually improves the accuracy of predictive model and profitability. Purpose: Finding the relevant technical indicators for financial market is crucial for financial market prediction. There is no particular set of technical indicators that fits all financial market since every market has its own dynamics and it is important to first select the technical indicators that reflects the market conditions rather than using the most commonly employed technical indicators. Design/Methodology/Approach: Random forest technique is used to find the relevant technical indicator. And each indicator is ranked according to Gini index. Findings: The analysis suggest that the set of technical indicators is not the same for every financial market. Further with in the categories of developed markets, emerging markets and frontier markets the set of technical indicators changes. Implications/Originality/Value: Hence, it is summarized that to improve the forecasting ability of the predictive model it is important to find the market specific technical indicators that will end up in higher profit for investors.

KEYWORDS: Technical Indicators, Random Forest, Feature Selection, Financial Forecasting

1. INTRODUCTION

Analysis of financial data is pertinent and arduous concern in economy. Most of the financial data is treated as time series. This type of data is associated with forecasting in literature. In recent years stock market forecasting has gained popularity. As financial data include numerous indicators, feature selection techniques are utilized to reduce the dimensionality of data and study it efficiently by reducing redundancy. Hence, feature selection is given a prime importance in various researches to get unbiased results. Feature selection techniques are not confined to financial data only. It’s been used in number of other fields like medicine, microarray technology, text-mining and economy. A mixed approach is used for selecting marker genes in microarray gene dataset (Wang, Makedon, Ford, & Pearlman, 2004). Further, a comparative analysis of ten most commonly used feature selection techniques was conducted on binary class biomedical data (Drotár, Gazda, & Směkal, 2015) and a new feature selection technique was developed for gene classification(Rakkeitiwina, Lursinsap, Aporntewan, & Mutirangura, 2015). Similarly, genetic algorithm, artificial neural network (Kim, 2006), a wrapper technique (C.-J. Huang, Yang, & Chuang, 2008) and a filter-based feature selection (C.-L. Huang & Tsai, 2009) were suggested for stock market dataset. Also, filter and wrapper approach hybrid (Crone & Kourentzes, 2010) was proposed for feature selection of time series data. Artificial neural networks (ANN) and support vector machines (SVM) for classification problems have been used for financial data (Vapnik, 2013).
Since, financial time series data is nonlinear and chaotic (Abu-Mostafa & Atiya, 1996), ANN can predict price movements accurately (Kimoto, Asakawa, Yoda, & Takeoka, 1990; Quah, 2006; Roman & Jameel, 1996; Tsaih, Hsu, & Lai, 1998). SVM can also predict future price behavior (Kim, 2003; Pai & Lin, 2005; Thissen, Van Brakel, De Weijer, Melssen, & Buydens, 2003). Therefore, soft computing technique can select features (Blum & Langley, 1997; Kim & Han, 2000; Liu & Motoda, 1998). But before utilizing any of the soft computing technique and datamining, data cleaning is a prerequisite. Other than handling the missing and duplicate values, standardization of data, it’s equally important to reduce its dimensionality by finding the relevant features and removing irrelevant features from the data (Guyon, Nikravesh, & Zadeh, 2006). Dimensionality reduction handles the problem of overfitting, improves model precision and curtail the “curse of dimensionality”. The input reduction can be done by feature selecting or extracting technique. Feature extraction is based on projections or compression by altering original features. However, feature selection preserves the original features and select the best features.

There is a vast literature on feature selection of financial market (Kara, Boyacioglu, & Baykan, 2011) selected ten indicators based on the knowledge and review of field expert and prior researches (Armano, Marchesi, & Murru, 2005; Diler, 2003; C.-L. Huang & Tsai, 2009; Kim & Han, 2000; Kumar & Thenmozhi, 2000; Yao, Tan, & Poh, 1999). Patel, Shah, Thakkar, and Kotecha (2015) also use 10 technical indicators following the footsteps of Kara et al., 2011 for studying the predictive accuracy of Artificial Neural Network (ANN), Support Vector Machine (SVM), random forest and naïve-Bayes. Evaluation is carried out on 10 years of historical data from 2003 to 2012 of two stocks namely Reliance Industries and Infosys Ltd. and two stock price indices CNX Nifty and S&P Bombay Stock Exchange (BSE) Sensex. Later Shynkevich, McGinnity, Coleman, Belatreche, and Li (2017) follow (Armano et al., 2005; Kara et al., 2011; Kim, 2003) for technical indicator selection. (de Oliveira, Nobre, & Zarate, 2013) utilize eight technical indicators as used by the qualified public to predict price behavior. Taş and Gürsoy (2016) employ four successful technical indicators for comparing their effectiveness with fuzzy indicators.

Berutich, López, Luna, and Quintana (2016) include some popular technical indicators along with other financial metrics like CAPM alpha and beta, and the Sharpe ratio to find the best trading strategy. Dash and Dash (2016) consider six popular technical indicators for a hybrid stock trading framework. Hsu eHsu, Lessmann, Sung, Ma, and Johnson (2016) employ seven technical indicators based on previous literature for testing predictive accuracy, which is higher if a model incorporates technical indicators. Yin and Yang (2016) use eighteen technical indicators as suggested by previous academic studies for oil price prediction. Ozturk, Toroslu, and Fidan (2016) use 24 popular technical indicators for developing heuristic trading rules for forex market. Weng, Ahmed, and Megahed (2017) consider three technical indicators for forecasting one day ahead price movements. Bruni (2017) use technical indicators for forecasting price trend based on previous work (Murphy, 1986). Basak, Kar, Saha, Khaimed, and Dey (2019) select six technical indicators to determine future direction of stock prices using tree-based classifier. But in all these studies choice of technical indicators is not been done through any feature selection technique rather based on prior studies or the popularity of technical indicator.

In this study, the extensive list of 90 technical indicators are analyzed using the nine stock market indexes covering advanced, emerging and frontier countries categories of Morgan Stanely Capital International (MSCI). The focus of this study is to find the relevant subset of technical indicators selected with random forest feature selection approach. In previous studies as discussed earlier, authors employed the most commonly used technical indicators without paying much attention to market specific indicators. The technical indicators which work well for advanced countries may not be good enough for frontier markets.

2. DATA AND METHODOLOGY

2.1 Data sample

This paper considers nine diversified financial markets from a group of developed, emerging and frontier markets of MSCI market index categorization. This sample support the experimental design of the study.
### Table 1: List of selected countries and stock markets

<table>
<thead>
<tr>
<th>Sr.</th>
<th>Countries</th>
<th>Ticker Symbols of Stock Market Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>United States</td>
<td>GSPC</td>
</tr>
<tr>
<td>2</td>
<td>United Kingdom</td>
<td>FTSE</td>
</tr>
<tr>
<td>3</td>
<td>France</td>
<td>FCHI</td>
</tr>
<tr>
<td>4</td>
<td>Pakistan</td>
<td>KSE</td>
</tr>
<tr>
<td>5</td>
<td>China</td>
<td>SSEC</td>
</tr>
<tr>
<td>6</td>
<td>India</td>
<td>NSEI</td>
</tr>
<tr>
<td>7</td>
<td>Sri Lanka</td>
<td>CSE</td>
</tr>
<tr>
<td>8</td>
<td>Argentina</td>
<td>MERV</td>
</tr>
<tr>
<td>9</td>
<td>Bangladesh</td>
<td>DSEX</td>
</tr>
</tbody>
</table>

#### 2.2 Random Forests

Breiman (2001) proposed random forest (RF) by combining the concept of random subspaces and bagging. Random forest grows number of uncorrelated decision tree to find the best classification output. Random forest takes the following steps in growing decision trees:

1. A random sample of various cases is considered. Later, with replacement sampling for other trees is done.
2. A subset of variables (m) is selected and based on Gini score best split is determined. High values of m increase the correlation of trees resultantly increase the predictive ability of tree by lowering the error rate. There is a wide range of values of m as this is the only parameter in the RF to which the model is sensitive.
3. Out of sample testing is done on 33% of the cases to compute prediction error rate.
4. Variable importance is computed on out-of-sample data down the grown trees, number of votes for the target variable are counted. Then each variable values are independently changed. Then to produce a measure of effect number of votes for the tree with different variables is deducted from the number of votes for the unaltered trees. Lastly, average effect is computed for all trees to show the variable importance value.

#### 3. EXPERIMENTAL DESIGN

![Experimental design diagram]

*Fig. 1: Experimental design*

The selected indexes data is retrieved from [www.finance.yahoo.com](http://www.finance.yahoo.com). The target variable is the price change computed as:

\[
\text{PriceChange} = -Cl(stock) - Op(Stock)
\]

\[
\text{Class} = \text{ifelse(PriceChange} > 0, "Up", "Down").
\]
Then technical indicators are calculated. All the variables are arranged in data frame. The data set is divided into 70% training dataset and 30% testing dataset. The ranking of variable is displayed in variable importance plot.

4. RESULTS AND DISCUSSION

We have nine different country indices covering developed, emerging and frontier categories of MSCI. In all these nine indexes authors employ 90 technical indicators. Random forest technique is used to find the most pertinent features. Our findings reveal that these TI indicators are not same for all nine indexes. The variable importance and selection vary with each market. As per our results Closing price 2 (closing price of last two days) is ranked highest in all indexes except in NSEI, FCHI and DSEX where it is replaced by CLV and Wilders respectively. The Gini Index score of these indicators also change for every index. Closing price 2 is score 1652.29 for GSPC, 449.87 for KSE, 580.45 for MERV, 557.42 for CSE, 407.36 for SSEC and 747.64 for FTSE. Authors have summarized the results by focusing on top five indicators as shown in Table 02. Closing 2, CLV, and Closing 5 are common in all selected indicators except DSEX however ranking is different.

![Fig. 2: Variable Importance of GSPC](image2)

![Fig. 3: Variable Importance of FTSE](image3)
Fig. 4: Variable Importance of FCHI

Fig. 5: Variable Importance of KSE

Fig. 6: Variable Importance of SSEC
Fig. 7: Variable Importance of NSEI

Fig. 8: Variable Importance of CSE

Fig. 9: Variable Importance of MERV
Table 2: Comparison of top five technical indicators

<table>
<thead>
<tr>
<th>Developed Markets</th>
<th>Emerging Markets</th>
<th>Frontier Markets</th>
</tr>
</thead>
<tbody>
<tr>
<td>GSPC Closing 2</td>
<td>FTSE Closing 2</td>
<td>MERV Closing 2</td>
</tr>
<tr>
<td>FCHI CLV</td>
<td>KSE Closing 2</td>
<td>DSEX Wilders</td>
</tr>
<tr>
<td>SSEC CLV</td>
<td>NSEI CLV</td>
<td></td>
</tr>
<tr>
<td>CSE CLV</td>
<td>CSE CLV</td>
<td></td>
</tr>
<tr>
<td>MERV CLV</td>
<td>DSEX Wilders</td>
<td></td>
</tr>
</tbody>
</table>

5. CONCLUSION

This paper investigates the relevance of technical indicators with different financial markets. Our results show that there is no set of technical indicators that work well for all markets. It’s a common practice to utilize most popular technical indicators for predicting stock prices (Shynkevich et.al, 2017; Bruni 2017, Basak et.al, 2018). But our findings negate it. The selection of TI and their ranking differs with the change of market. The sample used in this study is well diversified among advance, emerging and frontier markets but results does not support one size fits all. Even selection and ranking of TI varies in each category. Our findings are addition in current literature regarding TI selection which leads to improve predictive model accuracy and reduces the problem of overfitting.

REFERENCES


