

# Ranking of Stocks Using a Hybrid DS-Fuzzy System: An Application to Bombay Stock Exchange

Gour Sundar Mitra Thakur, Rupak Bhattacharyya,  
Seema Sarkar (Mondal) and Partha Kar

**Abstract** The most crucial phase in the construction of any stock portfolio is proper selection of stocks before the assignment of investment ratio. Presence of various factors which directly or indirectly influence the stock market, have made this selection task uncertain and challenging. In this paper Dempster-Shafer (DS) evidence theory and fuzzy rule based system have been hybridized to rank the stocks under Bombay Stock Exchange based on their historical performance. This ranking proved to be very fruitful for the selection of stocks for short-term investment period.

**Keywords** Ranking of stocks · Fuzzy rule based system · Critical factors · DS theory

---

G.S. Mitra Thakur (✉)

Department of Computer Science and Engineering, Dr. B. C. Roy Engg. College,  
Durgapur, West Bengal, India  
e-mail: cse.gsmt@gmail.com

R. Bhattacharyya

Department of Mathematics, Bijoy Krishna Girls' College, Howrah,  
West Bengal, India  
e-mail: mathsrup@gmail.com

S. Sarkar (Mondal)

Department of Mathematics, National Institute of Technology,  
Durgapur, West Bengal, India  
e-mail: seemasarkarmondal@yahoo.co.in

P. Kar

Department of Computer Science, Proper Web Technologies Pvt. Ltd.,  
Kolkata, West Bengal, India  
e-mail: partha.kar.diatm@gmail.com

© Springer Science+Business Media Singapore 2016

R.K. Choudhary et al. (eds.), *Advanced Computing and Communication Technologies*,  
Advances in Intelligent Systems and Computing 452,  
DOI 10.1007/978-981-10-1023-1\_11

## 1 Introduction

Portfolio selection is an optimal allocation of investors' asset among maximum number of stocks to provide good return with less risk. Predicting the performance and selecting the portfolio of stocks have been an important and complicated research area for researchers in modern stock exchanges because of uncertainties. Uncertainties come from various technical and fundamental factors which influence the stock market directly and indirectly.

In 1952, in a ground-breaking research by Markowitz [1], by representing portfolio return in the form of mean and risk as variance, he proposed a portfolio selection model to maximize return and minimize the risk. Then the relationship between risk and mean-variance were examined by many researcher like Zhou et al. [2], Yong et al. [3] etc. According to Modern Portfolio Theory (MPT), the investors always try to maximize their profit with minimum amount of risk by selecting a proper portfolio. Xidonas et al. [4], divided the portfolio selection process into two different stages. In the first stage some stocks are selected to construct the portfolio. In the second stage the percentage of the total value for each stock is identified.

In the last few decades Artificial Intelligence is playing a very vital role in the selection of stocks and in the construction of portfolio. Among various soft computing tools fuzzy set theory and fuzzy logic have been widely used in MPT. Details of few of these researches can be found in references like [5–8].

As the selection of proper stocks plays the most vital role in portfolio construction, main objective of this research is to identify top 10 stocks, which are expected to perform better for short-term investment period, out of 30 stocks registered in Bombay Stock Exchange (BSE) based on their historical performance. In this proposed work a new hybrid DS-fuzzy system has been developed for the ranking of stocks. Initially four critical factors namely Price to Earning Ratio (P/E), Price to Book Value (P/B), Price to Sales ratio (P/S) and Long Term Debt to Equity ratio (LTDER) have been identified which have direct or indirect influence in the performance of stocks. Then the historical data (2003–04 to 2011–12) of these four factors have been used to rank the stocks under BSE with the help of proposed DS-Fuzzy rule based system.

In Sect. 2 design of the proposed DS-Fuzzy model has been discussed and Sect. 3 draws the conclusion.

## 2 Design of the Proposed DS-Fuzzy Model for Stock Evaluation

Dempster-Shafer evidence theory and fuzzy rule based expert system have been hybridized in the proposed model to deal with the inherent uncertainty in stock selection problem. Dempster-Shafer evidence theory has been briefly introduced in

this section before the discussion of the proposed model. For detailed explanation regarding expert system and its various components reader can consult [9].

### 2.1 Dempster-Shafer Evidence Theory

A research by A. P. Dempster in 1966 [10], proposed a multivalued mapping from one space to another space. It has been used for statistical inference, when we have multiple sample information and we need to identify a single hypothesis. The DS evidence theory of belief function was first proposed by A. P. Dempster in 1968 and then extended by Shafer in 1976 [11] which can deal with incomplete and uncertain information.

The basic idea of Dempster-Shafer Theory mainly deals with four components: frame of discernment, basic probability assignment (BPA), plausibility function (*Pl*) and belief function (*Bel*). Frame of discernment is defined as a finite set of mutually exhaustive and exclusive hypotheses. Let  $\theta = \{\theta_1, \theta_2, \theta_3, \dots, \theta_n\}$  is the frame of discernment, where  $n$  is the total number of hypotheses.  $X$  is assumed to be an arbitrary subset of  $\theta$ . Then belief function over  $\theta$  can be expressed as a mapping function  $m$  defined as:

$$m : 2^\theta \rightarrow [0, 1] \tag{1}$$

Such that  $m(\phi) = 0$  and  $\sum_{X \subseteq \theta} m(X) = 1$ .  $m(X)$  is the value of basic probability assignment (BPA) for a given set  $X$ .

Dempster’s rule of combination for combining the degree of belief of two sets  $m_1$  and  $m_2$  can be defined as follows:

$$m_3(Z) = \frac{\sum_{X \cap Y = Z} m_1(X)m_2(Y)}{1 - \sum_{X \cap Y = \emptyset} m_1(X)m_2(Y)} \tag{2}$$

### 2.2 Proposed Model

The design of the proposed model is described through the following stages:

**Collection of Historical Data and Fuzzification** Total 30 stocks are registered in BSE. Last 9 years’ (2003–04 to 2011–12) data of four factors (P/E ratio, P/B value, P/S ratio and LTDER) have been collected from different web sites like [www.capitaline.com](http://www.capitaline.com), [www.bseindia.com](http://www.bseindia.com), [www.nseindia.com](http://www.nseindia.com). For the simplicity of this system we have normalized the data within the range of [0, 10] by considering the maximum historical value of the last 9 years data to be 10.

**Table 1** Membership functions for the linguistic values of the input variables

Factors in terms of linguistic variables	Linguistic values	Fuzzy trapezoidal membership
Price to earnings ratio (range 0–10)	Low	(0 0 1.8 2.8)
	Standard	(1.7 3.5 4.6 5.8)
	High	(5.3 7.5 10 10)

Fuzzification is a process of taking a crisp value as input and transforming it into the degree required by the terms. In the proposed model all four input factors have been converted to fuzzy variables with three linguistic values (Low, Standard and High) using fuzzy trapezoidal memberships in [0, 10]. Sample fuzzy membership for P/E is given in Table 1.

**Fuzzy Rule Construction** Dempster-Shafer theory has been used to construct fuzzy rules. Generally, the knowledge base of a Fuzzy Inference System is developed by using collection of fuzzy rules which determines how the output will be generated based on the given input. This proposed system consists of four input variables P/E, P/B, P/S and LTDER as discussed earlier. Based on these input parameters DS-fuzzy inference system determines whether the Selection of any stock will be *Highly Favorable*, *Moderately Favorable* or *Not Favorable* as an output.

As mentioned earlier in Table 1, as an example ‘Price to Earnings Ratio’ is having three linguistic values: *Low*, *Standard* and *High*. D-S theory has been applied to construct the fuzzy rules. The frame of discernment has been considered as  $\theta = \{High\ Performance, Average\ Performance\ and\ Poor\ Performance\}$  to represent the three possible outcomes for each stocks.

Though there are various other ways to measure the performance of stocks, one of the most commonly used measure is Risk/Return ratio. Risk is generally measured by semivariance (S) of the stocks’ previous returns and Return (R) is actually the mean of previous returns. Lower value of this ratio indicates better performance of stocks.

Now based on the linguistic value of each factor and comparing those with S/R value as a measure of their performance, basic probabilities have been assigned towards different hypotheses. The data of 2012–13 has been used to identify the stock. Initially a standard value for each of these four factors has been set for every stock with the help of expert opinions. From the last 9 years’ historical data it has been found that if the value of P/E ratio is around its standard value, then 75 % of stocks under BSE perform better and S/R value is also found to be satisfactory. So 0.75 degree of belief has been assigned towards the hypothesis *{High Performance}*. If the value of P/E ratio is much lower than its standard value then it has been found that 60 % of stocks under BSE perform average and S/R value is average. So 0.6 degree of belief has been assigned towards the hypothesis *{Average Performance}*. Similarly, when the value of P/E ratio is much higher than the standard value then it has been found that 70 % of stocks under BSE perform poor and S/R value is also unsatisfactory. So 0.7 degree of belief has been assigned

**Table 2** Membership functions for the linguistic values of the input variables

Factors	Linguistic values	High performance (H_P)	Average performance (A_P)	Poor performance (P_P)
P/E	Low		0.6	
	Standard	0.75		
	High			0.7
P/B	Low			0.8
	Standard		0.6	
	High	0.65		
P/S	Low			0.7
	Standard		0.65	
	High	0.75		
LTDER	Low	0.6		
	Standard		0.75	
	High			0.65

towards the hypothesis {*Poor Performance*}. In the same way initial believes have been assigned for all other factors (Table 2).

Let us consider a simple case

Rule 1. IF Price to Earnings Ratio is **Standard** THEN Performance will be **High** ( $m_1(H\_P) = 0.75$ ).

Rule 2. IF Price to Book Value is **Low** THEN Performance will be **Poor** ( $m_2(P\_P) = 0.8$ ).

Rule 3. IF Price to Sales Ratio is **High** THEN Performance will be **High** ( $m_4(H\_P) = 0.75$ ).

Rule 4. IF LTDER is **Low** THEN Performance will be **High** ( $m_6(H\_P) = 0.6$ )

Now Dempster Rule of combination has been applied to calculate the final mass value of the final rule.

Let the IF part of Rule 1 and Rule 2 be the first two evidences and  $m_1, m_2$  be two mass functions for them. Degree of belief has been assigned to the hypotheses based on the evidences. So from Rule 1,  $m_1(H\_P) = 0.75$  and  $m_1(\theta) = (1 - 0.75) = 0.25$ ,  $m_1(\theta)$  represents the degree of belief in the rest of the hypotheses present in the hypothesis set. And from Rule 2,  $m_2(P\_P) = 0.8$  and  $m_2(\theta) = (1 - 0.8) = 0.2$ ,  $m_2(\theta)$  represents the degree of belief in the rest of the hypotheses present in the hypothesis set.

Now these two evidences are combined and new mass  $m_3$  is generated as mentioned in Table 3 and mass value for hypothesis  $m_3$  can now be calculated with the help of Eq. (2).

**Table 3** Combination of mass considering first two evidences

Combining $m_1$ and $m_2$	$m_2(P\_P) = 0.8$	$m_2(\theta) = 0.2$
$m_1(H\_P) = 0.75$	$\emptyset = 0.6$	$H\_P = 0.15$
$m_1(\theta) = 0.25$	$P\_P = 0.6$	$\theta = 0.05$

**Table 4** Combination of mass considering first three evidences

Combining $m_3$ and $m_4$	$m_4(H\_P) = 0.75$	$m_4(\theta) = 0.25$
$m_3(H\_P) = 0.375$	$H\_P = 0.2813$	$H\_P = 0.09$
$m_3(P\_P) = 0.5$	$\emptyset = 0.375$	$P\_P = 0.125$
$m_3(\theta) = 0.125$	$H\_P = 0.0934$	$\theta = 0.03125$

**Table 5** Combination of mass considering first three evidences

Combining $m_5$ and $m_6$	$m_6(H\_P) = 0.6$	$m_6(\theta) = 0.4$
$m_5(H\_P) = 0.7499$	$H\_P = 0.4499$	$H\_P = 0.2999$
$m_5(P\_P) = 0.2$	$\emptyset = 0.12$	$P\_P = 0.08$
$m_5(\theta) = 0.05$	$H\_P = 0.03$	$\theta = 0.02$

Now consider the IF part of the rule 3 to be the new evidence and  $m_4$  be the mass function. A belief of 0.75 has been assigned towards the hypothesis  $\{High\ Performance\}$ .

So from Rule 3,  $m_4(H\_P) = 0.75$  and  $m_4(\theta) = (1 - 0.75) = 0.25$ . Again  $m_3$  and  $m_4$  are combined to generate new mass  $m_5$  as mentioned in Table 4 and new mass values for hypothesis  $m_5$  can now be calculated with the help of Eq. (2).

Finally, the IF part of the Rule 4 is the last evidence and  $m_6$  is the mass function. A belief of 0.6 has been assigned towards the hypothesis  $\{High\ Performance\}$ .

From Rule 4,  $m_6(H\_P) = 0.6$  and  $m_6(\theta) = (1 - 0.6) = 0.4$ . Now,  $m_5$  and  $m_6$  are combined and final mass value  $m_7$  is generated as mentioned in Table 5 and final mass value for hypothesis  $m_7$  can be calculated using Eq. (2) as shown in Eq. (3).

$$m_7(H\_P) = \frac{0.4499 + 0.2999 + 0.03}{1 - 0.12} = 0.8861$$

$$m_7(P\_P) = \frac{0.08}{1 - 0.12} = 0.09 \tag{3}$$

$$m_7(\theta) = \frac{0.02}{1 - 0.12} = 0.023$$

In this way D-S theory is applied and final mass values for rest of the 80 rules are calculated. Maximum final mass value for *High Performance* ( $H\_P$ ) is found to be 0.9916 and minimum final mass value for *High Performance* ( $H\_P$ ) is found to be 0. The conclusion part of each rule is decided on the basis of their final mass values for *High Performance* ( $H\_P$ ). Now, favorability of the stocks is divided into three categories: *Highly Favorable* (Final mass value for  $H\_P$  between 0.76 and 1), *Moderately Favorable* (Final mass value for  $H\_P$  between 0.46 and 0.75) and *Not Favorable* (Final mass value for  $H\_P$  less than 0.45). So the combined version of the above sample rules becomes:

“IF Price to Earnings Ratio is **Standard** AND Price to Book Value is **Low** AND Price to Sales Ratio is **High** AND LTDER is **Low** THEN the stock is **Highly Favorable**”.

Because the final mass value for *High Performance (H\_P)* was 0.8861 (between 0.76 and 1). Finally total of 81 (3<sup>4</sup>) rules are formulated for the knowledge base of the proposed DS-fuzzy inference system and above three selection categories were converted into fuzzy linguistic variables with trapezoidal membership values for the output of the DS-fuzzy inference system as shown in Table 6.

**Ranking of Different Stocks Under BSE** Data of 2012–13 for all four input factors of 30 stocks are used as input to the DS-fuzzy system and then they are ranked based on their defuzzified values in descending order. This model has identified the top 10 stocks out of 30 stocks as short term investment preference. Table 7 shows the details of top 10 Stocks based on this ranking. As the consequent of every rule indicates about the favorability of stocks higher defuzzified value indicate higher favorability. Among all the 30 stocks, the highest defuzzified value derived as 0.8644 corresponds to *Hindustan Unilever Ltd.* and it has topped the ranking and the lowest defuzzified value is obtained for *Tata Steel Ltd.* as 0.1626.

As mentioned earlier S/R ratio is used as performance measure of stocks very often, a ranking of stocks is also done based on the S/R values of the stocks for the FY 2013–14. When we compared this ranking with our proposed ranking we found ten matches in top 15. It gives a clear indication about the efficiency of the model.

**Table 6** Fuzzy trapezoidal membership of the output variable

Output variable	Linguistic values	Fuzzy trapezoidal membership
Selection (range 0–1)	Not favorable	(0 0 0.172 0.448)
	Moderately favorable	(0.34 0.46 0.57 0.75)
	Highly favorable	(0.64 0.88 1 1)

**Table 7** Top 10 stocks

Rank	Name of the stocks	Defuzzified values (input data: FY 2012–13)
1	Hindustan Unilever Ltd.	0.8644
2	Sun Pharmaceutical Inds. Ltd	0.8457
3	I T C Ltd.	0.5369
4	Coal India Ltd.	0.5340
5	Tata Consultancy Services Ltd.	0.4936
6	Infosys Ltd.	0.3842
7	Dr. Reddy’s Laboratories Ltd.	0.3553
8	Bajaj Auto Ltd.	0.3292
9	Hero Motocorp Ltd.	0.3002
10	Cipla Ltd.	0.2850

So it is obvious that any investor preparing a portfolio based on this proposed ranking for short-term investment period is expected to get better return in FY 2013–14.

### 3 Conclusion

Selection of suitable stocks from any stock market to maximize return is always a tough and challenging task for any kind of investor. This paper proposes a model to rank stocks under BSE based on their historical performance. A novel hybrid DS-Fuzzy model is proposed for this purpose. DS theory is used to assign beliefs towards different stocks and decide the consequents of fuzzy rule base and finally stocks are ranked based on the fuzzy rule based system. As the tested result shows a satisfactory comparison with the recent performance of the stocks this model can be implemented for any stock exchange. The ranking of this model can be easily used in any rank preference based portfolio selection model.

### References

1. Markowitz, H.: Portfolio selection. *J. Finance* **7**, 77–91 (1952)
2. Zhou, X.Y., Li, D.: Continuous-time mean-variance portfolio selection: A stochastic lq framework. *Appl. Math. Optim.* **42**(1), 19–33 (2000)
3. Fang, Y., Wang, S.: An interval semi-absolute deviation model for portfolio selection. In: *Fuzzy Systems and Knowledge Discovery*, pp. 766–775. Springer (2006)
4. Xidonas, P., Ergazakis, E., Ergazakis, K., Metaxiotis, K., Askounis, D., Mavrotas, G., Psarras, J.: On the selection of equity securities: an expert systems methodology and an application on the athens stock exchange. *Expert Syst. Appl.* **36**(9), 11966–11980 (2009)
5. Hiemstra, Y.: A stock market forecasting support system based on fuzzy logic. In: *Proceedings of the Twenty-Seventh Hawaii International Conference on System Sciences*, 1994, vol. 3, pp. 281–287. IEEE (1994)
6. Bermudez, J.D., Segura, J.V., Vercher, E.: A fuzzy ranking strategy for portfolio selection applied to the Spanish stock market. In: *Fuzzy Systems Conference, 2007. FUZZ-IEEE 2007*, pp. 1–4 (2007)
7. Fasanghari, M., Montazer, G.A.: Design and implementation of fuzzy expert system for Tehran stock exchange portfolio recommendation. *Expert Syst. Appl.* **37**(9), 6138–6147 (2010)
8. Nair, B.B., Dharini, N.M., Mohandas, V.: A stock market trend prediction system using a hybrid decision tree-neuro-fuzzy system. In: *2010 International Conference on Advances in Recent Technologies in Communication and Computing (ARTCom)*, IEEE, pp. 381–385 (2010)
9. Giarratano, J.C., Riley, G.D.: *Expert systems: principles and programming*. CENGAGE Learning (2005)
10. Dempster, A.P.: Upper and lower probabilities induced by a multivalued mapping. *The annals of mathematical statistics*, pp. 325–339 (1967)
11. Shafer, G., et al.: *A mathematical theory of evidence*, vol. 1. Princeton University Press, Princeton (1976)